Bazele teoretice ale evaluarii psihologice- Rezumat

Acest rezumat se bazează pe patru materiale

Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului,* Ed. Fundatiei Romania de Maine, Bucuresti,
Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition),*Routledge, London and New York,
Urbina Susana, *2004- Essential of psychological testing*, John Wiley & Sons, Inc. Hoboken, New Jersey,

Cuprins

1. LOCUL ŞI ROLUL PSIHODIAGNOZEI ÎN CADRUL ŞTIINŢEI ŞI PRACTICII PSIHOLOGICE; FUNCŢIILE PSIHODIAGNOZEI

1.5. Delimitări conceptuale

1.5.1. Psihodiagnoza şi evaluare psihologică
1.5.2. Concepte cheie: examen psihologic, instrument psihologic, observaţia în examenul psihologic, conduita examinatorului
1.5.3. Procesul decizional în evaluarea psihologică
1.5.4. Psihodiagnoza formativă- tendinţa contemporană

1.1. DELIMITĂRI CONCEPTUALE

1.1.1. PSIHODIAGNOZĂ ŞI EVALUARE PSIHOLOGICĂ

Evaluarea psihologică este un set de proceduri utilizate de psiholog cu scopuri care includ, în general:

1. Producerea de imagini (evaluarea ca suport pentru a dezvolta descrieri şi imagini despre o anumită persoană, privind funcţiunile intelectuale, trăsăturile de personalitate, aptitudinile speciale, dificultăţile de învăţare, variabile emoţionale şi sociale).
2. Producerea de decizii (evaluarea persoanei are ca finalitate producerea de decizii despre relaţia acelei persoane cu mediul ei actual sau cu mediile posibile: adecvare, normalitate, functionalitate, disfuncţionalitate);
3. Verificarea de ipoteze, construirea unei teorii (evaluarea ca suport pentru testarea ipotezelor despre personalitate).

Psihodiagnoza este activitatea specifică ce foloseşte diferite tipuri de instrumente pentru a obţine informaţii despre structura, dinamica psihică şi personalitatea unei persoane. Astfel de instrumente utilizate sunt: testul psihologic; alte metode precum observaţia comportamentului verbal sau non-verbal, interviul.

După unii autori în ceea ce priveşte delimitările între evaluarea psihologică şi psihodiagnoză nu sunt consacrate şi clare în literatura de specialitate. Spre exemplu evaluarea psihologică este considerată a avea mai degrabă o dominată calitativă, în timp ce psihodiagnoza este predominat cantitativă(Mitrofan, N. 2007).
În ceea ce urmează vom delimita accepţiunea clasică a psihodiagnozei de cea modernă.

Criteriile psihodiagnosticului traditional mai ales sub aspectul psihometriei clasice se referă la:

- natura standardizată a testelor aptitudinale
- intervenţia minimală a examinatorului astfel incat subiectul sa recurga numai la propriile resurse intelectuale
- sistemul de evaluare se opreste la nivelul comunicarii rezultatelor ca atare, intr+un mod artificial
- examinarea sub forma unui interviu in care subiectul nu este informat despre pertinenţa răspunsurilor sale
- eventualele încurajări premise de standardizare au o scăzută autenticitate, pentru că nu implică valoarea răspunsului ci urmăresc doar obţinerea unei atitudini positive a subiectului faţă de situaţia de testare
- condiţiile testării sunt identice pentru toţi subiecţii, deşi nu toţi sunt pregătiţi pentru a se adapta la acestea(unii se acomodează mai greu cu interacţiunea limitativă şi nu pot profita autentic de acest tip de relaţie cu psihologul); unii dintre subiecţi au o eficienţă semnificativ crescută dacă primesc întăriri positive şi/sau repetate, în timp ce alţii nu sunt influenţaţi de astfel de conduite

Evoluţii în psihodiagnoza modernă

In anii 60 ai secolului XX a apărut curentul cunoscut sub numele de **evaluarea comportamentală**, care încearcă să determine ce anume fac persoanele, unde, cînd şi în ce circumstanţe.
În evaluarea şi predicţia comportamentului se porneşte în primul rînd de la comportamentul motor şi verbal manifest şi se renunţă la interpretarea rezultatelor la teste. Evaluarea de tip comportamental se delimitează de evaluarea psihometrică atît în privinţa scopului cît şi în privinţa metodelor şi principiilor implicate(Fernandez- Ballesteros, 1993).
Principiile evaluarii comportamentale sunt:

- Conceperea observării ca metodă principală de măsurare
- Abandonarea chestionarelor de personalitate- acestea fiind considerate incapabile să descrie probabilitatea răspunsurilor comportamentale sau variabilele care afecteaza aceste raspunsuri
- Renunţarea la testele tradiţionale care evaluează trăsături psihice precum structurile mentale
- Renunţarea la principiile psihometrice ale testelor

Abordări prezente ale evaluării comportamentale subliniază necesitatea de a se lucra şi în noul cadru teoretic cu instrumente corespunzătoare ca fidelitate şi validitate, verificate psihometric, deşi iniţial curentul a pornit cu negarea psihometriei şi a necesităţilor metrologice serioase. Astfel continuă să fie aplicate testele psihologice, ca si chestionarele de personalitate, mai ales cele care privesc factorii de influenţa ai comportamentului ce nu pot fi observaţi direct, precum motivaţia.

- Au fost create metode şi instrumente proprii de măsurare:
  - Tehnici de observare
  - Proceduri fiziologice
  - Interviuri structurate
  - Tipuri noi de scale de evaluare de către alte persoane

Analizînd modelele utilizate de evaluarea comportamentală, D.J. Cone, în 1986 (Silva, 1993) găseşte doar două mari categorii, care le subsumează pe toate celelalte, şi anume abordarea nomotetică a trăsăturii, şi abordarea idiografică a comportamentului.
Abordarea nomotetică a trasăturii- are ca obiect de studiu trăsături sau sindroame(cum ar fi anxietatea socială, depresia, agorafobia), şi dezvoltă instrumente pentru deducerea acestora. În acest model instrumentul este adecvat dacă scorurile între indivizi variază, adică dacă diferenţiază între indivizi privind aceeaşi trăsătură. Ideea de baza este că orice trăsătură este universal aplicabilă tuturor persoanelor şi în consecinţă se încearcă predicţia comportamentului prin măsurarea trasăturilor. Capacitatea de predicţie mică a acestor instrumente provine din realitatea că fiecare persoană are trasături relevante care-i influenţează comportamentul. Dacă trasăturile măsurate de test nu sunt relevante atunci nu poate fi facută o predicţie bună a comportamentului(Paunonen, Jackson, 1985). În abordarea nomotetică se caută legile generale sau ecuaţiile care se aplică tuturor persoanelor. Procedura statistică preferată este analiza de regresie, adică identificarea funcţiei de regresie care leagă predictorul – trăsătura de personalitate- de criteriu- comportamentul specific. Pentru obţinerea ecuaţiei de regresie se utilizează date culese pe un eşantion, iar ecuaţia serveşte la predicţia comportamentului tuturor indivizilor din populaţia din care a fost selectat eşantionul. Abordarea nomotetică a trăsăturii defineşte de fapt concepţia care stă la baza evaluării psihodiagnostice tradiţionale(Silva, 1993).
Abordarea idiografică a comportamentului- studiază un comportament specific şi dezvoltă procedee pentru evaluarea sa inductivă. O procedură este considerată adecvată ţinînd cont de variaţia scorurilor la acelaşi individ. Ideea de bază este că fiecărei persoane îi sunt specifice anumite determinante ale comportamentului şi pe acestea încearcă să le identifice, sub forma ecuaţiilor proprii fiecărui individ. În acest context teoretic şi metodologic se identifică persoanele pentru care este potrivită o anumită ecuaţie de predicţie a comportamentului pe baza trăsăturii.

Deosebirea principală dintre evaluarea psihometrică şi cea comportamentală constă din presupunerea asupra comportamentului unei persoane.
În teoria psihometrică scorul obţinut la un test este considerat a fi compus dintr-un scor real şi o eroare care provine din măsurare. Scorul real reflectă o entitate internă stabilă, deci nu se schimbă, cel puţin într-un interval scurt de timp, iar dacă apare o modificare de scor la test aceasta provine din eroarea de

măsurare. Un test este considerat bun dacă scorurile obţinute de o persoană la măsurări succesive nu diferă prea mult. Această calitate a testului se numeşte fidelitate. Comportamentul este considerat a fi dependent în primul rînd de o entitate internă stabilă- trăsătura.
În concepţia evaluării comportamentale, comportamentul nu este, în mod necesar stabil. Dacă scorurile unei persoane la acelaşi test diferă se consideră că aceasta se poate datora chiar schimbărilor de comportament, nu neapărat impreciziei instrumentului de măsură. Cînd apelează la teste, evaluarea comportamentală caută să determine variabilele responsabile şi nu face nici o presupunere asupra erorii de măsurare. Deoarece susţine dependenţa de situaţie a comportamentului evaluarea comportamentală nu utilizează validarea concurentă ca mijloc de predicţie şi nici nu pretinde valori mari ale coeficientului de fidelitate test-retest.

1.1.2. CONCEPTE CHEIE: EXAMEN PSIHOLOGIC, INSTRUMENT PSIHOLOGIC, OBSERVAŢIA ÎN EXAMENUL PSIHOLOGIC, CONDUITA EXAMINATORULUI

Examen psihologic
Orice examen psihologic implică următoarele aspecte comune:

- Relaţia între examinator şi examinat(subiectul examinării)
- Evaluarea rezultatelor subiectului de către examinator
- Comportamentul examinatorului
- Calităţile specifice ale sarcinii

Deşi este o interrelaţie între două persoane în care ambele caută informaţii, oferă şi primesc feedback, scopul specific al testării este de a informa examinatorul. Acesta va genera o serie de ipoteze despre modul cum funcţionează subiectul într-un număr de situaţii sau interacţiuni diferite. Factorii care pot influenţa sistematic informaţiile sunt clasificaţi în doă categorii: factori care ţin de metodologia de testare(metoda de testare, secvenţa de prezentare a testului sau testelor) şi factori care ţin de examinator(personalitate, trebuinţe, caracteristici psihofizice). Printre variabilele studiate şi descrise în literature de specialitate referitor la examinator se regăsesc:

- Caracteristicile fizice ale examinatorului
- Prezenţa-absenţa acestuia în sala de testare
- Comportamentul său”cald” sau “rece”
- Condiţionarea operată asupra comportamentului verbal al subiectului
- Efectul examinatorului ca persoană în afara evaluării personalităţii sale
- Efectul examinatorului ca persoană cu evaluarea personalităţii sale

Subiectul reacţionează nu numai la stimulii pe care-i reprezită testul ci şi la toţi ceilalţi stimuli care-l înconjoară, deci la stimuli relevanţi pentru testare dar şi la stimuli nerelevanţi. Produsele subiectului sunt astfel un rezultat al reacţiilor sale particulare la itemii testului şi la examinator, ca întreg. Administrarea testelor de către examinatori diferiţi în aceeaşi situaţie, sau de către acelaşi examinator dar în situaţii diferite poate conduce la rezultate diferite. Printre variabilele care afectează implicarea subiectului regăsim setul subiectiv, clasa socială a subiectului şi stresul situaţional. O altă sursă de eroare, legată de posibilitatea interpretării eronate, este constituită tocmai de faptul că examinatorul trebuie să ia decizii privind răspunsurile subiectului. Examinatorul poate interpreta eronat datorită percepţiilor sale subiective, caracteristicilor sale de personalitate sau chiar cunoaşterii limitate a valorii stimulului de către subiect. Studiile au constatat tendinţa examinatorilor de a examina diferit protocoalele testelor proiective şi în funcţie de clasa socială a subiectului (pentru protocoale în esenţă similare cei din clasa de jos au primit interpretări mai nefavorabile decât cei din clasa de mijloc).

Forme ale examenului psihologic
Cel mai important criteriu, obiectivul şi gradul de formalizare, clasifică examenele psihologice în două extreme:

1. Examenul psihometric
2. Examenul clinic

Examenul psihometric are principii şi metodologie foarte formalizată, şi care pornesc de la definiţia lui E.L. Thorndike: "Dacă un lucru există, xistenţa sa are o anumită măsură cantitativă". În consecinţa psihologia se ocupă cu trăsături distincte, care au o existenţa reală, şi pe care le posedă toţi oamenii în măsuri diferite. Examenul psihometric caută să măsoare aceste trăsături reale.

Examenul clinic se deosebeşte în primul rând printr-un accent mult scăzut pe aspectul formal şi mult mai mare pe problemele caracteristice, specifice celui examinat.
Atât în privinţa calităţilor psihometrice, a standardizării cât şi a gradului de deschidere faţa de specificul subiectului examinat foarte multe aspecte ale instrumentelor psihologice se leagă de dihotomozarea examinării în formală şi deschisă.
Totuşi există multe forme mixte datorită diversităţii scopurilor şi situaţiilor care cer examinare.
Cea mai formalizată situaţie dintre toate este reprezentată de examenul de selecţie, în privinţa condiţiilor de examinare, a administrării probelor, a interpretării acestora. Fiecare probă este inclusă într-o baterie şi are o anumită pondere, precis calculată, în diagnoza sau predicţia performanţei în anumite tipuri de cerinţe profesionale.
În alte situaţii, spre exemplu consiliere educaţională, de cuplu, de carieră- probele sunt înalt standardizate, însă interpretarea este mai centrată pe nevoile clientului, deci mai puţin standardizată. Pot fi incluse şi probe mai puţin standardizate, iar interviul este mai puţin formalizat.

Instrument psihologic
Un instrument psihologic este o procedură sistematică de a evalua comportamentul uneia sau mai multor persoane. Caracteristicile profesionale ale testului ca instrument de evaluare sunt obiectivitatea şi standardizarea.

Standardizarea

Un instrument de măsură profesionist trebuie să evalueze într-un mod precis, standardizat, pe o scală de unităţi egale. Standardizarea se referă şi la procedura de lucru, procedura de înregistrare a rezultatelor, cotarea rezultatelor şi interpretarea acestora. Orice condiţii care afectează performanţa trebuie specificate în manualul testului.

Obiectivitatea

Obiectivitatea înseamnă înregistrarea observaţiilor pentru a elimina erorile de rememorare, notarea răspunsurilor după anumite reguli. Cele mai obiective instrumente sunt testele cu răspunsuri corecte pentru că se utilizează o grilă de corectare stabilă. Se spune despre un test că este obiectiv dacă el permite să se măsoare fără ambiguitate capacităţile unei persoane (Klausnitzer, 1992). În literatura de specialitate se menţionează trei condiţii ale obiectivităţii:

- Rezultatele să nu depindă de nivelul pregătirii subiectului în domeniul testării psihologice
- Fiecare răspuns la item să fie interpretat la fel de către mai mulţi experţi
- Scorul să nu fie influenţat de personalitatea examinatorului, adică să aibe cotaţie obiectivă

Ultima dintre condiţii nu este îndeplinită de multe teste, în special cele proiective şi cele de cunoştinţe cu răspuns deschis. Cu toate acestea se poate ajunge la un grad de obiectivitate şi în aceste situaţii. Deşi testul nu este pe deplin obiectiv, certitudinea interpertărilor reprezintă un aspect al obiectivităţii acestuia.
Tipuri de instrumente psihologice
Din perspectiva gradului de standardizare se pot regăsi aceleaşi două extreme: probe foarte formalizate- denumite în general- obiective, şi probe deschise către problemele specifice individului, cu un grad mai mic de formalizare, ca ambiguitate a testului şi variabilităţii răspunsurilor- printre care se numără probele proiective.

Observaţia în examenul psihologic

În orice examen psihologic este necesară completarea rezultatelor cu alte date. Observaţia este utilizată pentru completarea cu date obiective a rezultatelor şi desprinderea semnificaţiei lor complete. Şi actul observaţiei poate fi făcut în două maniere polare- una strict obiectivă, formală, neutră, după o grilă, şi o altă manieră, mai puţin analitică, dar focalizată pe imaginea globală a persoanei examinate. Dezavantajul observaţiei sistematice, obiective este dificultatea de a integra datele obţinute într-o imagine globală, cu sens, a persoanei. În practică cele două maniere se pot completa. Observarea are ca obiecte, de regulă, atitudinea subiectului faţă de examen, faţă de examinator, variaţiile în comportamentul rezolutiv al subiectului, încrederea în sine, interesul pentru teste.

Conduita examinatorului

Atitudinea celui care examinează poate evolua între neutru, obiectiv şi personal, subiectiv. În funcţie de caracteristicile examenului şi ale subiectului examinat atitudinea se poate adapta. În ipostaza formală, examinatorul are un comportament standard, oferă explicaţii sau instrucţiuni standard, neimplicativ, înregistrează răspunsurile sec, evită abordarea personală, în scopul de a nu influenţa în nici un fel situaţia, şi de a păstra standard situaţia de examen pentru toţi subiecţii. Atitudinea personală, subiectivă şi mai puţin formală, deschisă maschează caracterul de examen al situaţiei şi este utilă în relaţie cu subiecţii anxioşi. În condiţiile examenului clasic, psihometric subiectul nu trebuie ajutat când este în dificultate, în afara situaţiilor când este evidentă suferinţa cauzată de nereuşită. Explicaţiile sunt scurte şi la obiect. În situaţia evalurii dinamice, formative subiectul este încurajat, fără a-i fi oferite informaţii legate de răspunsuri. Supravegherea în timpul examenului nu permite perturbarea spontaneităţii comportamentului şi activităţii subiectului, pentru a nu-i creea insecuritate, mai puţin în cazurile evidente de fraudă.

În literatura de specialitate se menţionează unele trăsături de dorit ale examinatorului:

- Sănătate şi securitate emoţională
- Simţul umorului
- Flexibilitate
- Abilitatea de contact
- Focalizarea pe rezolvarea probleme prin intermediul aplicării psihodiagnozei- nu doar pentru constatarea performanţei în termeni cantitativi- clasificare
- Cu competenţa teoretică compatibilă cu explicarea comportamentului concret şi înţelegerea fundamentului teoretic al testului
- Cu competenţa privind selectarea celei mai adecvate, unitare şi mai integrate baterii de teste pentru o situaţie concretă

1.1.3. PROCESUL DECIZIONAL ÎN EVALUAREA PSIHOLOGICĂ

Justificarea esenţială pentru testare este de a obţine informaţii care să sprijine rezolvarea unei probleme prezente. Datorită faptului că pe baza rezultatelor se poate interveni asupra vieţii unei persoane este nevoie să fie controlate cât mai multe variabile posibil iar strategia să fie cea mai adecvată. Intervenţia este eficientă în măsura înţelegerii naturii problemei şi astfel prin utilizarea unor instrumente adecvate, formale şi informale pentru a obţine o informaţie bogată şi variată, relevantă pentru decizie.

Aceste condiţionări impun un prim set de reguli:

- Testarea psihologică se face doar când există o problemă specifică, suficient de bine definită pentru a permite decizia privind adecvarea testării.
- Testarea impune necesitatea ca examinatorul să dispună de capacitatea de a traduce informaţia dată de test în prescripţii cu sens, practice şi adecvate situaţiei şi problemei ce se cere rezolvată, inclusiv la nivel de recomandare practică pentru cel examinat.
- Testarea impune cunoaşterea şi folosirea adecvată a consideraţiilor etice şi deontologice privind utilizarea testelor psihologice.

Modelări ale strategiei de decizie
Există două niveluri de decizie în sarcina psihodiagnostică
1. clasic- clasificarea euristică
2. prin modelare- evaluarea comparativă

Prima abordare reflectă modul tradiţional de concepere a diagnosticului în medicină, respectiv se bazează pe "sisteme nozologice"(o colecţie cunoscută de informaţii cheie ce pot fi identificate prin simptome observabile) şi un set informal de principii care ghidează procesul de inferenţă, exprimat sub forma unor judecăţi de tipul- "dacă…..atunci…" şi raportează rezultatul la un diagnostic.

A doua abordare se bazează pe un model structural al sistemului şi pe un model formal de comportament, definibil ca "normal", "echilibrat", "adecvat", versus "inadecvat", "deficient", "patologic".
Psihodiagnosticianul trebuie să decidă prin identificarea unui set de cauze sau condiţii care indică o anumită funcţionare pozitivă sau negativă. Psihodiagnosticul este o formă particulară de explicaţie a funcţionării pozitive sau negative cu sprijinul unei teorii asupra psihicului. În modelele de tip constructiv se procedează prin recunoaşterea simptomelor şi raportarea lor la un model de sistem. Psihodiagnoza constructivă poate construi astfel toate tipurile de disfuncţii care pot apare în interiorul unui sistem, chiar şi pe cele nemanifestate încă.

Fenomenologia testării psihologice
Conceptul de fenomenologie a testării a fost formulat de Snygg şi Combs, în 1959, şi reprezintă abilitatea de a surprinde şi de a înţelege cadrul intern de referinţă al examinatului. Datorită multiplelor influenţe ce se manifestă în testare înţelegerea cadrului subiectului, din punctul său de vedere este importantă pentru că permite evaluarea cu sens asupra sa(diagnostică şi prognostică) şi cu sens pentru acesta, şi permite găsirea unei modalităţi de rezolvare a unor probleme. Înţelegerea cadrului intern de referinţă al subiectului implică sensibilitate, intuiţie şi atenţie în special privind comportamentul non-verbal, nu numai verbal, şi poate fi dificilă pentru examinatorul rigid, defensiv. Examinatorii rigizi şi defensivi se simt ameninţaţi şi induc astfel anxietate şi subiectivism în relaţia cu cel examinat.
Decizia poate fi influenţată de trei mari zone informaţionale: mediul exterior- sociocultural şi fizic; psihicul subiectului- cu nivele diferite de accesibilitate; şi instrumentele psihologice- cu nivele diferite de profunzime.
Mediul ţine de caracteristicile concrete ale situaţiilor de testare şi de stările mediului fizic.
Psihicul subiectului prezintă un nivel accesibil observaţiei şi un nivel mediat multiplu- infrastructura psihică, substratul neuro-biologic şi substratul genetic. Psihismul subiectului este un tot unitar care reprezintă o unitate funcţională a mai multor instanţe:
- Una vizibilă, observabilă, comportamentală
- Conştiinţa- structura psihică conştientă- despre care subiectul poate relata
- Substratul inconştient al psihismului- infrastructura psihică profundă- pe care subiectul nu o cunoaşte şi nu o poate relata
- Dincolo de instanţele psihice există o infrastructură funcţională, biologică- neuro-biologică şi genetică

Cei mai semnificativi factori ce pot afecta starea subiectului sunt:
- Anxietatea
- Nevoia de a se prezenta într-o lumină favorabilă
- Nesiguranţa în răspunsuri
- Setul de expectaţii mentale
- Deprinderile cognitive
- Prejudecăţile şi idiosincrasiile
- Abilităţile verbale şi de exprimare, fluenţa verbală şi de raţionament
- Capacitatea de înţelegere prin lectură

Dintre toţi factorii anxietatea se manifestă în toate tipurile de teste, dar mai ales în cele obiective, de performanţă şi în interviu, deoarece subiectul se află în situaţia existenţei unui posibil pericol la adresa stimei de sine. Anxietatea este acel factor general de tensiune şi instabilitate care face ca persoana să nu se simtă în largul ei, neliniştită de teama eşecului, a necunoscutului.Alte surse ale anxietăţii sunt necunoaşterea modul de reacţie la teste, relaţia cu personalitatea şi conduita examinatorului, seturi de expectaţie şi reacţie neconforme cu ceea ce se aşteaptă de la subiect. Aceasta poate afecta răspunsurile subiectului, de aceea este recomandată clarificarea acestuia privind modul de desfăşurare a testării şi liniştirea acestuia. Un exerciţiu posibil pentru reducerea anxietăţii are trei secvenţe: ieşirea(imaginativă) din situaţia imediată, imaginarea unei alte situaţii de perfectă securitate şi linişte, readucerea în situaţie.
Din perspectiva instrumentelor utilizate acestea pot viza performanţa, autoevaluarea, evaluarea celorlalţi, evaluarea unor dimensiuni ale personalităţii prin metode psihometrice sau prin metode proiective. Acestea aduc informaţii din diverse unghiuri asupra subiectului:

- Din unghiul capacităţilor şi aptitudinilor subiectului, ce ţin de comportament şi de conştiinţă- adresându-se nivelului conştient
- Din unghiul imaginii de sine- adresându-se conştientizării de sine
- Din unghiul probelor nestructurate, în care subiectul proiectează predominant conţinuturi inconştiente- probele proiective (ambiguitatea sarcinii accesează nivelul inconştient; de exemplu testul Szondi accesează nivelul biologic şi genetic al unităţii care este omul).
- Din unghiul auto-evaluării cât şi din unghiul hetero-evaluarii(evaluarea celorlalţi, sau a examinatorului). Evaluarea celorlalţi se produce tot la nivel conştient, vizibil, dar aduce în plus propriile elemente inconştiente(idiosincrasii şi prejudecăţi, reperul evaluărilor fiind imaginea de sine a evaluatorului).

1.1.4. PSIHODIAGNOZA FORMATIVĂ- TENDINŢĂ CONTEMPORANĂ

Deoarece este o tendinţa tot mai actuală, mai ales în evaluarea aptitudinală, psihodiagnoza formativă merită o abordare separată. Diagnosticul, în acest tip de abordare, este formativ(nu constatativ) pentru că poate ghida intervenţia, în sensul evidenţierii funcţiilor psihice şi operaţiilor gândirii, deficitare şi care necesită dezvoltare, pentru creşterea performanţelor. Diagnosticul pune accentul, nu atât pe constatarea performanţei şi pe această baza a clasificării şi a predicţiei performanţei viitoare, cât pe capacitatea persoanei de a se schimba, în urma unei etape de învăţare. Astfel ceea ce se măsoară este un potenţial de învăţare, nu un coeficient intelectual. În evaluarea tradiţională psihometrică indivizii sunt comparaţi raportându-se la normele specifice vârstei, spre exemplu în ceea ce priveşte inteligenţa academică. Aceasta se bazează pe presupoziţia că că ei au avut oportunităţi egale de a învăţa anterior testării. Principala inovaţie pe care evaluarea formativă o aduce în câmpul psihometriei este legată de semnificaţia acordată coeficientului de inteligenţă. Deşi acesta reprezintă o măsură a performanţei într-un anumit moment, într-un anumit context, privind capacitatea de a profita şi învăţa din experienţele trecute la care a fost expus un subiect, există puţine motive să asumăm sau să acceptăm că performanţa la un test de inteligenţă furnizează o măsură stabilă şi de încredere a performanţei viitoare. Valorile mici ale coeficienţilor de validitate de criteriu(predictivă) demonstrează acest lucru. Evaluarea formativă pune accentul pe capacitatea de schimbare a fiinţei umane şi influenţa factorilor de mediu, chiar şi în cazurile când există afecţiuni organice, limitând astfel influenţa factorilor genetici, generatori de discriminare socială, din perspectiva autorilor acestei abordări(Feuerstein, R., 1979).
Sursa dificultăţilor cognitive ale subiectului poate fi localizată într-una din cele trei faze ale actului mental: inputul, elaborarea, outputul. Majoritatea pot fi determinate de ceea ce se cheamă deficit al medierii în învăţare, în experienţa de viaţă a subiectului. Ca atare intervenţia de îmbogăţire instrumentală a subiectului trebuie să se bazeze pe medierea învăţării.
În timpul evaluării examinatorul poate susţine procesul de rezolvare a problemelor prin trei tipuri de intervenţii de ajutor:

- Sunt utilizate sarcini a căror rezolvare nu depinde în mod necesar de achiziţiile cognitive anterioare;
- Sunt anihilate acele obstacole actuale, prezente, care duc la ineficienţa funcţionării cognitive;
- Este indus un nivel optim de funcţionare mentală, în funcţie de natura sarcinii şi de capacitatea subiectului.

Interacţiunea examinator- subiect nu este o simplă repetare, amplificare de informaţii sau reluare a unor modele habituale ale subiectului, ci vizează îmbogăţirea repertoriului de interacţiuni. Conceptul teoretic de la care s-a pornit este acela de "zonă proximă a dezvoltării", concept introdus de doi psihologi ruşi Luria şi Vîgotski. Zona proximă a dezvoltării reprezintă o acea zonă virtuală în care o capacitate actuală poate evolua. Principiul fundamental este reducerea distanţei dintre performanţă şi competenţă, şi crearea unor condiţii şi contexte ale situaţiei problematice care pot facilita şi valorifica potenţialele latente. În cadrul examenului au loc mai multe secvenţe:

- Faza de aplicare iniţială, în care se măsoară capacitpţile rezolutive ale subiectului, fără a se acorda ajutorul examinatorului, prin itemi din teste psihometrice tradiţionale(ex. Raven)
- Faza de învăţare în care dacă subiectul a greşit rezolvarea la anumiţi itemi i se oferă ajutorul pentru rezolvarea corectă sau pentru rezolvarea celor nerezolvaţi, fiind orientat în aşa fel încât să descopere singur- învăţare mediată
- Faza de evaluare a gradului de transferabilitate şi de integrare în structurile cognitive, operatorii ale subiectului, a cunoştiinţelor introduse în faza a doua. Sunt utilizaţi itemi ai unor teste paralele.
- Se utilizează tipuri de probleme adaptate nivelului optim al eficienţei funcţiilor mentale, în funcţie de vârstă sau de tipul de deficienţă mentală
- Se ameliorează componentele metacognitive- cele care ţin de motivaţie atenţie, ce intervin în funcţionarea cognitivă.

Scorurile obţinute în faza 1 și în faza 3 sunt utilizate la determinarea profitului cognitiv, acesta fiind o masura mai valida a potenţialului de învăţare, şi un predictor mai bun al performanţei viitoare.
Rezultatele se referă la patru aspecte:

1. Măsurarea validă a competenţelor cognitive ale subiectului
2. Determinarea dificultăţiilor funcţionale majore care apar în timpul rezolvării sarcinii
3. Compararea nivelului de dezvoltare cognitivă cu potenţialul de învăţare
4. Delimitarea condiţiilor psiho-pedagogice susceptibile pentru activarea potenţialelor cognitive

Evaluarea dinamică, formativă presupune schimbări metodologice în următoarele patru domenii:

1. **În structura testelor**- Instrumentele de testare trebuie să fie construite într-un mod care să ofere examinatorului şi examinatului sarcini ce pot fi utilizate într-un proces de învăţare şi care pot permite examinatorului să evalueze efectul procesului de învăţare asupra capacităţii individului de a face faţă unor situaţii noi.
2. **În situaţia de examinare**- Trecerea de la un scop static la unul dinamic, implică o schimbare în situaţia de testare care va transforma examinatorul într-un profesor-observator şi examinatul într-un student-executant. Această transformare atrage după sine o varietate de schimbări în interacţiunea obişnuită şi stabilirea unui proces de comunicare în dublu sens.
3. **În orientarea testelor**- schimbarea de accent de la rezultate la procesul de obţinere a rezultatelor este o altă schimbare necesară pentru a trece de la o abordare statică la una dinamică.
4. **Interpretarea rezultatelor**- Vârfurile de performanţă între rezultatele obţinute trebuie utilizate ca o indicaţie a potenţialului examinatului.

Unul dintre cei mai importanţi reprezentanţi ai acestei orientări este Reuven Feuerstein.

Principala contribuţie a noii direcţii de dezvoltare poate fi rezumată prin accentul pus pe evaluarea intelectului prin sarcini de învăţare. În cele ce urmează pe parcursul întregului manual ne vom situa în cadrul psihometric, al psihometriei clasice în cea mai mare parte.

1.2. FUNCŢIILE PSIHODIAGNOZEI

Psihodiagnoza ca activitate sistematică de evaluare psihologică are următoarele funcţii:

1) Surprinderea corectă a trăsăturilor şi capacităţilor psihice individuale şi evidenţierea variabilităţii psiho-comportamentale intra-grupală versus grup de referinţă
2) Evidenţierea cauzelor care au condus spre o anume realitate prezentă, în special în cazul disfuncţiilor sau tulburărilor psihice
3) Prognoza evoluţiei probabile a comportamentului persoanei în anumite condiţii şi situaţii contextuale, în funcţie de repere
4) Evidenţierea cazurilor de abatere în sens pozitiv sau negativ de la o normă(etalon) de dezvoltare psiho-comportamentală; se pune problema distincţiei între normalitate şi anormalitate
5) Evidenţierea sau validarea(demonstrarea funcţionalităţii) programelor de învăţare şi formare
6) Formarea unor capacităţi de cunoaştere şi autocunoaştere
7) Sprijinirea unor decizii în consiliere şi orientare vocaţională (evidenţierea de capacităţi şi motivaţie)
8) Sprijinul deciziilor de conduită în demersul din psihoterapie, asistenţă, consultanţă psihologică
9) Verificarea unor ipoteze ştiinţifice, prin utilizarea testelor în cadrul experimentelor ştiinţifice

1.3. PRUDENŢĂ ŞI ABUZ ÎN UTILIZAREA TESTELOR PSIHOLOGICE

1.3.1. CONSECINŢE NEGATIVE ALE TESTĂRII PSIHOLOGICE

Pe măsură ce testarea psihologică s-a afirmat ca un curent principal în evaluarea psihologică, în toate domeniile aplicative au apărut tot mai evidente unele pericole implicate. Cele mai semnificative se referă la condiţia persoanei examinate. Astfel este importantă conştientizarea următoarelor aspecte:

- Testarea psihologică reprezintă o invadare a vieţii private, a intimităţii unei persoane. De aceea testarea psihologică se face numai cu consimţământul informat al persoanei examinate. Atunci când este vorba de un minor, dreptul la intimitate trebuie exercitat de părinţii acestuia.
- Evaluarea poate duce la creşterea anxietăţii persoanei, ceea ce poate conduce la efecte adverse. Printre acestea enumerăm influenţa asupra comportamentului şi a performanţei la test, tulburări nedorite în viaţa de familie, consecinţe incorecte sau invalidante în plan social ale interpretărilor rezultatelor. Evaluarea trebuie, deci, să ţină cont de efectele anxietăţii.
- Tendinţa spre autoîmplinire a celui care examinează sau care interpretează, dacă nu este suficient conştientizată, poate constitui o sursă de interpretare şi diagnoză greşită a rezultatelor testului.
- Interpretarea greşită la o etichetare incorectă. Etichetarea incorectă a unei persoane va conduce spre perceperea socială predominant din perspectiva etichetei. Consecinţele pot fi anxietatea şi jena în familie, şi, mai grav, limitarea oportunităţilor celui etichetat de a-şi dezvolta adecvat potenţialul său general.
- Oricât de complexă ar fi, o singură probă nu este suficientă pentru o decizie, de aceea este necesară o baterie de teste, sau un set de probe, adecvate scopului urmărit.

- Utilizarea neprofesionistă a informaţiilor confidenţiale rezultate la un test sau interviu poate duce la consecinţe grave pentru persoana testată.
- Este necesar ca persoana care examinează să fie matură şi securizată emoţional, deoarece poate influenţa reacţia subiectului la test şi rezultatele acestuia.

Problemele care se au în vedere în comunitatea ştiinţifică, referitor la reducerea consecinţelor negative ale testării sunt legate de :

- Impactul negativ asupra minorităţilor şi femeilor
- Descurajarea persoanelor testate şi "instituţionalizarea" eşecului pentru indivizi
- Predarea pentru testare şi limitarea curriculum-ului şi învăţării
- Accentul pus pe sarcinile de memorare în detrimentul celor care implică procese cognitive mai complexe, necesare pentru succesul în lumea reală
- Crearea de bariere sociale şi organizaţionale
- Direcţionarea şi specializarea prematură, limitând posibilităţile de schimbare şi variaţie

1.3.2. NEVOIA DE PSIHODIAGNOZĂ

Consecinţele pozitive ale utilizării adecvate a evaluărilor validate sunt legate chiar de scopurile măsurării psihologice:

- Sunt un ghid pentru luarea deciziilor- selecţia celor mai potriviţi candidaţi, promovarea bazată pe performanţă, premierea bazată pe performanţă
- Permit o eficienţă şi o eficacitate mare- decizii bazate pe volum mare de informaţii relevante, în timp scurt
- Permit controlul calităţii
- Au efect în protejarea publicului care beneficiază de serviciile companiilor- prin reducerea angajărilor neglijente în ocupaţii critice
- Permit obiectivitatea pentru a realiza comparaţii şi decizii ce privesc indivizii, sau faţă de criterii obiective prestabilite
- Implică eficienţă în privinţa costurilor şi utilizării

Concluzie

O trăsătură comună a tuturor testelor este precizia lor limitată. Toate îmbunătăţirile în tehnologia testelor(evaluarea formativă, evaluarea comportamentală, creşterea fidelităţii şi validităţii, modele teoretice noi) urmăresc gestionarea şi reducerea acestei imprecizii. Sunt extrem de rare cazurile când un test furnizează o măsură exactă a unei variabile(reale) care are efecte importante asupra comportamentului uman. Din acest motiv, nici deciziile care se iau pe baza lor nu sunt întotdeauna corecte (falşii pozitivi- eroare de tip alfa, falşii negativi- eroare de tip beta). De exemplu, unii psihologi susţin că, din cauza lipsei de consistenţă a comportamentului de la o situaţie la alta, coeficienţii de corelaţie liniară dintre scorurile unui test de personalitate şi un comportament pe care acesta îl prezice nu depaşesc valoarea de 0,3 sau 0,4 (W. Mischel şi R.E. Nisbett, după Funder, 1983, apud. Albu M., 1998). Ca lucrurile să fie şi mai complicate apar şi consecinţe negative în urma abuzului, utilizării neetice a testelor, sau de către neprofesionişti.

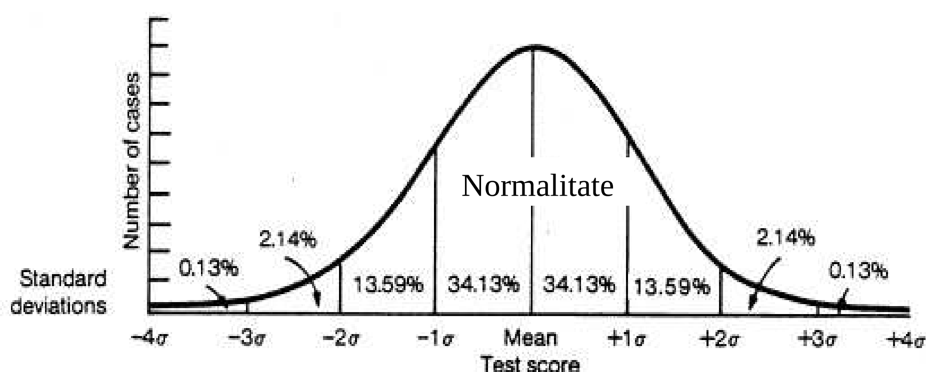
Dar, nu există alte metode de măsurare, ştiinţifice, mai precise. Prin urmare, testele psihologice reprezintă cea mai bună, cea mai corectă şi mai precisă tehnologie ştiinţifică disponibilă pentru a lua decizii importante asupra indivizilor (Murphy, Davidshofer, 1991, apud Albu, 1998).

1.7. CONCEPTUL DE NORMALITATE PSIHICĂ

Conceptul de normalitate psihică este folosit deseori în contextul psihologiei. De aceea este util să fie cunoscute cât mai corect accepţiunile sale. În mod curent funcţionează trei perspective asupra termenului de normalitate: perspectiva statistică, perspectiva social-culturală, perspectiva psihomedicală.

Perspectiva statistică
Se referă la normalitate ca regularitate caracteristică sau acceptată de un număr semnificativ de persoane. Are deci un suport cantitativ- norma- foarte pe larg spus: ceea ce este mai prezent într-o populaţie de referinţă este normal, ceea ce nu este prezent nu este normal, este în afara normalităţii. În psihologie accepţiunile cele mai utilizate sunt nivelul de excelenţă şi un model sau un etalon în raport cu care s-ar putea stabili o măsură sau formula o judecată(Lisievici P., 1997, apud Minulescu, 2003). În psihodiagnoză, delimitările dintre normalitatea psihică şi ceea ce iese în afara normalităţii se bazează pe masurători cantitative, dar şi calitative, cel mai adesea îmbinate. Variabilele la care se referă masurătorile sunt: comportamentul persoanei, cogniţiile, sentimentele sau situaţiile, nivelul de funcţionare sau destructurare al unor mecanisme şi abilităţi psihice, scopurile tratamentului, dar şi concepte teoretice precum eul, stima de sine, starea de bine, calitatea vieţii etc. Procesul de măsurare tebuie să fie precis, adică să aibă fidelitate şi validitate. Pentru ca scorurile să poată fi interpretate, să aibă sens în raport cu comportamentul uman în genere, se normează gama de rezultate posibile prin procedee statistice, adică se urmăreşte modul de distribuţie al rezultatelor de-a lungul unui continuum, şi modul în care se comportă populaţia de referinţă, respectiv aflarea tendinţei centrale- media, spre exemplu şi abaterea de la medie- abaterea standard a mediei, adică variabilitatea comportamentelor în populaţia respectivă.
Multe dintre variabilele psihologice se supun distribuţiei normale, descrisă matematic prin “curba lui Gauss”. Pe această curbă valorile medii ale variabilei reprezintă comportamentul majorităţii populaţiei(68,28%). Abaterile de la zona medie sunt egal distribuite. Abaterile moderate de la medie- cu valori cuprinse între o abatere standard de la medie şi două abateri standard- cuprind câte 13,59 % din populaţie(27,18%).
Abaterile extreme,cu valori cuprinse între două abateri standard şi trei abateri standard, reprezintă câte 2,14% din populaţie(4,28%). Abaterile reziduale reprezintă 0,13%(în total 0,26%).
Atunci când măsurăm cantitativ comportamentul unei persoane îl putem încadra în:

- Comportamentul majoritaţii- cei 68,28% din zona medie- şi atunci are un comportament normal (la fel ca majoritatea) pentru populaţia respectivă, cum se poate observa în figura de mai jos.
- Comportamentul unei minorităţi- cei 27, 18 %, din zonele imediat următoare
- Comportamentul unei excepţii semnificative de la medie, în plus sau minus- cei 4,53% din extreme



Pentru fiecare instrument şi fiecare condiţie de folosire a instrumentului se construiesc astfel de etaloane după criterii bine definite privind reprezentativitatea populaţiei.
În această accepţiune normalitate înseamnă determinarea nivelului de eficienţă al unei abilităţi psihice sau a unei dimensiuni de personalitate în raport cu modul în care se comportă acestea în populaţia de referinţă. În cele două situaţii semnificaţiile unui scor se interpretează diferit.
Când este vorba despre abilităţi sau capacităţi psihice, spre exemplu nivelul inteligenţei, dacă scorul subiectului(care reprezintă performanţa sa), se află în zona corespunzătoare mediei, atunci îl considerăm în normalitate. Dacă abaterea este în plus, peste medie rezultatul este calificat drept "bun", dacă de află între **+1σ şi +2 σ** (litera grecească sigma reprezintă notaţia pentru abaterea standard a scorurilor de la medie), şi "foarte bun" dacă se află între**+2 σ şi +3σ.** Invers, rezultatul este calificat drept "slab" dacă se află între **-1σ şi -2 σ, şi "**foarte slab", dacă se află între **-2 σ şi -3σ.** Pentru extreme(+3 **σ** ,+4 **σ; -**3 **σ,-**4**σ**) vorbim de excepţii pozitive şi excepţii negative. În termeni statistici ieşirea în afara normalităţii poate însemna o performanţă de excepţie, dar şi un nivel de retard mental.
Când este vorba de dimensiuni, caracteristici sau factori de personalitate, în măsura în care aceştia se comportă parametric, gradienţii de comportament se raportează tot la curba lui Gauss. Astfel o dimensiune precum extraversie-introversie poate prezenta, în urma răspunsului subiectului, un scor la medie, în zona normalităţii(m _+ 1 **σ**). Din punct de vedere psihologic este calificat drept "ambiversie", ceea ce nu are sens pentru că nu indică o entitate de sine stătătoare, ci o nedefinire. Dacă subiectul are scoruri, peste medie cu mai mult de 1 **σ,** atunci putem vorbi de un comportament extravert normal, iar peste 3**σ,** de un comportament extravert extrem. Un comportament extrem poate aduce inconveniente pentru că indică, în acest caz, dependenţa de mediu şi evenimente, un grad de instabilitate ridicat şi inconstanţă afectivă în legăturile cu ceilalţi. La fel se interpretează şi scorurile care cad sub medie, dar în sensul introversiei, nu al anormalităţii. Deci nu întotdeauna extremele pot fi considerate ca ieşiri în afara normalităţii.

Perspectiva social-culturală
În acest context vorbim de norme sociale şi culturale, respectiv de reguli sociale sau juridice instaurate de societate pentru a favoriza relaţionarea nondistructivă între membrii societăţii, şi de modele de comportament, obiceiuri, cutume, prejudecăţi ca standarde scrise sau nescrise. Evaluarea curentă implică astfel de norme în construirea instrumentelor şi în modul cum subiecţii răspund la chestionare. Un comportament "normal" într-o anumită zonă geografică sau comunitate, poate fi socotit "anormal", deviant în spaţiul unei alte culturi sau subculturi din aceeaşi cultură.
Normele sociale şi juridice impun limite. Încălcarea acestora este sancţionată, până la măsura izolării sociale. În acest cadru discutăm de concepte precum **adaptare, comportament adaptativ, comportament de risc, comportament dezadaptativ sau antisocial, comportament prosocial**. Individul care ajunge în mod constant să sfideze legea sau norma socială, care nu se poate adapta, şi care poate comite infracţiuni repetitive sau bizare, are structurări de personalitate particulare, grupate sub denumirea de "tulburări de personalitate sociopate".

Perspectiva psihomedicală
În această perspectivă discutăm despre criteriul comportamentului auto sau heterodistructiv, şi despre condiţia de boală psihică, şi implicit despre conceptul de sănătate mintală. Există diferite grade de anormalitate psihică. O condiţie de anormalitate relativă poate fi considerată spre exemplu depresia, când persoana are tendinţa de izolare, cu pericolul latent de ideaţie distructivă sau suicidară, sau suspiciozitatea şi interpretativitatea, specifice structurilor paranoide, care afectează şi viaţa celor din jurul persoanei. Vorbim astfel de **personalităţi accentuate**, de **comportamente reactive la stres** sau la condiţii existenţiale extreme. Pentru astfel de situaţii există consilierea psihologică şi psihoterapiile de diferite orientări.
Pentru marea patologie psihică, în care boala domină viaţa persoanei, există psihiatria şi psihoterapia *intra muros.*

Alte concepte care intervin în problematica normalităţii în psihologie
Tot în contextul psihopatologiei discutăm de conceptul de **anormalitate**. Anormalitatea vizează mai ales comportamentul şi conduita. Anormalitatea apare ca o perturbare de fundal iar boala psihică implică un aspect procesual; anormalitatea se referă mai ales la structura şi organizarea psihică, iar boala implică un proces morbid. Anormalitatea se relevă mai ales în tulburările de comportament, iar boala ca proces morbid interesează patologia clinică.
Tulburarea psihocomportamentală, anormalitatea se manifestă ca discordanţă între dorinţa persoanei, acţiunea sa şi ceea ce se aşteaptă şi se cere de la aceasta în contextul social.
Criterii uzuale ale devianţei psihopatologice:

- Disconfort psihic foarte intens şi de durată mare(stare cronică de anxietate sau depresie)
- Ineficienţă cognitivă- incapacitatea de a gândi clar realitatea sau de îndeplinire a obligaţiilor sociale, profesionale, de familie- eşecul
- Tulburări în funcţionarea organismului
- Devianţa comportamentală de la normele sociale

Sănătatea mentală
Intervine în delimitarea conceptului de normalitate şi este definită, de obicei, prin următoarele caracteristici:

- Capacitatea de conştientizare, acceptare şi corectitudine în modul cum se concepe pe sine
- Stăpânirea mediului şi adecvarea în modul de a face faţă cerinţelor vieţii
- Integrarea şi unitatea personalităţii
- Autonomia şi încrederea în sine
- Percepere realistă şi încredere în sine
- Continuitatea dezvoltării personale spre auto-actualizare

Alte criterii ale normalităţii(după Diamond şi Ellis, apud Dobranici, L. 2007)

- constiinţă clară a eului personal
- capacitate de orientare în viaţă
- nivel înalt de toleranţă la frustrare
- autoacceptare
- flexibilitate în gândire şi acţiunea
- realism şi gândire antiutopică
- asumarea responsabilităţii pentru tulburările sale emoţionale
- angajarea în activităţi creatoare
- angajarea moderată şi prudentă în activităţi riscante
- conştiinţă clară a interesului social
- gândire realistă
- acceptarea incertitudinii şi capacitatea de ajustare a acesteia
- îmbinarea plăcerilor imediate cu cele de perspectivă

Maturitate psihică
Conceptul de maturitate psihică nu are o relevanţă directă pentru psihometrie dar prilejuieşte cele mai interesante şi fructuoase dezbateri, cu reverberaţii mai largi decât cadrul psihometriei, până la magnitudinea celor despre condiţia umană.

Prezentăm spre exemplificare câteva concepţii psihanalitice despre normalitate (L. Dobranici, 2008):
S.Freud: „Normalitatea este o ficţiune ideală; fiecare ego este psihotic într-un anumit moment într-o măsură mai mare sau mai mică”;
K.Eissler: „Normalitatea absolută nu poate fi obţinută, deoarece persoana normală trebuie să fie pe deplin conştientă de gândurile şi sentimentele sale”;

M.Klein : „Normalitatea este caracterizată prin tărie de caracter, capacitatea de a face faţă emoţiilor conflictuale, capacitatea de a trăi plăcerea fără a provoca conflicte şi capacitatea de a iubi”;
E.Erikson: „Normalitatea este capacitatea de a fi stapân pe perioadele vieţii: încredere/neîncredere; autonomie/îndoială; iniţiativă/vinovăţie; activitate, producţie/inferioritate; identitate/confuzie de rol; creaţie/stagnare; integritatea ego-ului/disperare”.
H.Hartmann: „Funcţiile libere de conflicte ale ego-ului reprezintă potenţialul persoanei pentru normalitate; măsura în care ego-ul se poate adapta la realitate şi poate să fie autonom sunt asociate sănătăţii mintale”;
K.Menninger : «Normalitatea este capacitatea de a se adapta lumii exterioare cu mulţumire şi cu capacitatea de a stăpâni fenomenul de aculturaţie”;
A.Adler: „Capacitatea persoanei de a dezvolta sentimente sociale şi de a fi productiv/creativ sunt legate de sănătatea mintală; capacitatea de a munci creşte stima de sine şi face persoana capabilă de a se adapta”.
R.E.Money-Kryle: “Normalitatea este capacitatea de a atinge deplina conştiinţă de sine care de fapt nu este niciodată pe deplin obţinută”.
O.Rank : « Normalitatea este capacitatea de a trăi fără teamă, vinovăţie sau anxietate şi aceea de a avea responsabilitatea propriilor acţiuni”.

După cum se observă abordările par destul de diferite, deşi sunt toate în acelaşi cadru teoretic, al psihanalizei. Prima constatare ar fi că nu există o coerenţă printre autori care fac parte chiar din aceeaşi orientare teoretică. Acelaşi lucru îl putem spune şi despre conceptul de maturitate. Nu există un model coerent unic în psihologie, aplicabil la orice condiţii, prin care să delimităm maturizarea.

În psihologie, utilizarea termenului de maturitate psihică impune criterii de judecată care se referă la structura, funcţionarea şi echilibrul psihic al persoanei. Cadrul cel mai general prin care putem defini normalitatea psihică, din perspectiva psihologiei umaniste, implică trei aspecte simultan:

- Capacitatea de a stăpâni activ mediul
- Manifestarea unei unităţi a personalităţii
- Capacitatea de a percepe corect lumea şi propria persoană

Aceste aspecte sunt câştiguri ale persoanei de-a lungul dezvoltării sale ontogenetice. O să înfăţişăm pe scurt una dintre viziuni, cea a personologului (de orientare psihanalitică) Erik Erickson despre competenţele care se câştigă la diferite vârste.

Încă din primele stadii copilul depinde foarte mult de oportunităţile care i se oferă. De la naştere până la 18 luni copilul rezolvă problema încrederii fundamentale (conflictul încredere-neîncredere) în urma căruia capătă speranţă. În copilăria timpurie, între 18 luni şi trei ani copilul rezolvă problema autonomiei(conflictul autonomie- ruşine), în urma căreia capătă auto-control, curaj şi voinţă. Între trei şi cinci ani copilul rezolvă conflictul iniţiativă- vină, în urma căuia începe să-şi propună şi să atingă propriile sale scopuri

Astfel vârsta şcolară presupune formarea unui sentiment de competenţă personală, dobândirea unui statut şi a unor deprinderi adecvate, acestea fiind criterii de evaluare ale etapei de vârstă. Perioada adolescenţei are de rezolvat problema identităţii, ceea ce duce la constituirea autonomiei şi a increderii în sine.
Perioada adultului tânăr ar trebui să conducă la rezolvarea crizei identităţii, în sensul capacităţii de a empatiza, a stabili relaţii apropiate cu semenii, a iubi. În perioada adultă maturitate înseamnă rezolvarea crizei caracteristice stagnării, în sensul întăririi capacităţii generative a persoanei, a formării structurilor atitudinal-afective care permit comportamentele de asumare a responsabilităţii, a deciziei şi independenţa în evaluare.
Perioada senectuţii înseamnă dobândirea înţelepciunii prin depăşirea crizei disperării şi conservarea integrării emoţionale. Putem utiliza aceste criterii pentru a stabili în ce măsură o persoană este adecvată vârstei sale, şi astfel matură, sau normală.

Individuare şi autoactualizare

Conceptul privind sensul personal al vieţii, care, faţă de psihanaliza freudiană, dă consistenţă şi înţeles dezvoltării individuale spre autonomie şi împlinire, a fost dezvoltat de Carl Gustav Jung, încă de la începutul secolului trecut. Pentru Jung, individuarea are sensul unei diferenţieri a tuturor structurilor psihicului printr-o continuă şi concomitentă integrare în sfera vieţii conştiente, tendinţă pusă în joc de funcţia transcendentă, de individuare, înscrisă în fiinţa umană. La Jung individuarea implică descoperirea conştientă şi intrarea în relaţie cu un conţinut obiectiv al psihicului, pe care, pornind de la filosofia indiană(Upanişade) şi bazându-se în cele din urmă pe spiritualitatea occidentală, îl denumeşte Sine. Astfel, în procesul de auto-actualizare a acestor conţinuturi obiective se dezvoltă o identitate a eului mai stabilă şi continuă, transformarea eului obişnuit în eul înţelept. În context, boala psihică reprezintă o gamă de simptome care semnalează ieşirea din această direcţie de creştere, ceea ce conduce la regresii şi stagnări nevrotice.

Ulterior, şcoala umanistă din psihologia americană(Abraham Maslow, Carl Rogers, Erich Fromm etc.) a lansat conceptul de auto-actualizare sau auto-realizare, ca un prag în care motivaţia firească de creştere, este depaşită la nivel uman printr-o meta-motivaţie care dă naştere acţiunii meta-trebuinţelor. Aceasta reflectă integrarea în sensul existenţial personal a unor valori ideale precum frumuseţea, unicitatea, ordinea, perfecţiunea, participarea. Abraham Maslow propune un set de indicatori care caracterizează o persoană ce a atins acest prag al dezvoltării umane. Aceste persoane constituie mai degrabă o excepţie fericită de la normă şi pot fi caracterizate prin:

- Limpezime în modul de percepere a realităţii şi capacitate de a accepta ambiguităţile existenţei
- Acceptare de sine şi a celorlalţi, fără judecată, şi un grad minim de anxietate su vinovăţie legată de sine
- Comportament spontan şi fantezie în gândire, dar nu sunt total neconvenţionali
- Centrarea pe o problematică, mai degrabă decât centrare pe sine
- Capacitate de a privi viaţa cu obiectivitate şi căutarea solitudinii
- Comportament independent, autonom, dar nu rebel
- Iubire a vieţii
- Trăiri de extaz, chiar legate de sentimentul de putere, cu momente de mare creativitate, de stăpânire a necunoscutului
- Implicare socială şi identificare simpatetică cu orice om
- Experienţe interpersonale profunde, dar, de regulă, cu puţini oameni
- Respect pentru fiinţă
- Cunoaşterea diferenţei dintre mijloc şi scop şi lipsa neliniştii că trebuie să suporte un anumit mijloc pentru a atinge un scop
- Un sens filosofic al umorului, spontaneităţii, jocului şi lipsa de agresivitate faţă de ceilalţi
- Capacitatea de a descoperi probleme şi sensuri, mai ales în domeniul pe care sunt centrate
- Faptul că nu permit "culturii" să le controleze.

Toate aceste discuţii despre normalitate permit psihologului să identifice mai clar în ce cadru trebuie să se plaseze atunci când utilizează rezultatele pe care o persoană le-a obţinut la un test psihologic.

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition),*Routledge, London and New York,
Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului,* Ed. Fundatiei Romania de Maine, Bucuresti,

Mitrofan Niculae, 2001- Psihometria şi direcţiile ei de dezvoltare la început de mileniu. În Zlate M. (coord.), *Psihologia la răspântia mileniilor*. Iaşi: Polirom.

Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare*, Editura Polirom, Iaşi

Urbina Susana, *2004- Essential of psychological testing*, John Wiley & Sons, Inc. Hoboken, New Jersey,

COLEGIUL PSIHOLOGILOR DIN ROMÂNIA (2005 şi 2006), *Acte normative*.

2. PERSPECTIVĂ ISTORICĂ ASUPRA PSIHODIAGNOZEI

2.1.Repere privind preocupările dinaintea construirii testelor
2.2.Evoluţia psihodiagnozei în secolul XX
2.3.Continuitate şi discontinuitate în evoluţia psihodiagnosticului în România

2.1. REPERE PRIVIND PREOCUPĂRILE DINAINTEA CONSTRUIRII TESTELOR

350 î.e.n. Aristotel afirmă importanţa observării obiective în cunoaşterea comportamentului uman şi formulează primele trei principii care pot să explice asociaţia de idei.
1650 Rene Descartes consideră relaţia corp - minte ca interacţiune.
1822. F. Bessel realizează primele experimente semnificative în măsurarea diferenţelor individuale în ce priveşte timpul de reacţie, cu aplicaţie la observaţiile astronomice.
1869. Studiul ştiinţific al diferenţelor individuale începe cu cercetările lui Sir F. Galton, văr cu Ch. Darwin, care studiază diferenţele individuale şi aplică ideea lui Darwin privind adaptarea selectivă la cerinţele mediului. începe măsurarea sistematică a diferenţelor umane privind înălţimea, mărimea capului şi viteza de calcul. în mare măsură, munca sa a fost orientată spre investigarea inteligenţei superioare şi abilităţilor, pe care le considera în mare măsură înnăscute. Contribuţia sa cuprinde atât măsurători fizice şi psihice, cât şi dezvoltarea unor tehnici de statistică prin care urmărea determinarea relaţiei dintre aceste măsurători.
In 1960 publică *Clasificarea oamenilor în funcţie de darurile lor naturale*, iniţiind astfel studiul ştiinţific al diferenţelor individuale. In 1884 prezintă la Londra un laborator antropometric, comparabil cu laboratorul german de psihologie experimentală al lui Wundt.
1879. Galton dă prima descriere sistematică a unui test de asociaţie liberă, iar în 1983 publică *Inquiries into human faculty and its developments,* care, după unii autori (M. Reuchlin), marchează trecerea de la metodele de laborator de psihologie experimentală la metode simple şi aplicabile pe eşantioane largi de subiecţi.
1884. S. Hull fondează primul laborator american, Universitatea J.Hopkins.
1888. J. Mckeen Cattell inaugurează un laborator de testare la Universitatea Pennsylvania. în 1890 introduce termenul de „teste mentale". Inventează „deviaţia standard" ea măsură statistică a variabilităţii rezultatelor în raport de medie.
1891. Munsterberg propune în Germania o serie de probe pentru Studierea diferitelor profesiuni.
1896. Kraepelin raportează testarea comparativă a celor bolnavi mental cu oamenii normali. Studii experimentale privind memoria.
1898. Expediţia Torres Straits introduce pentru prima oară studiul Indivizilor analfabeţi cu ajutorul unor instrumente psihologice (McDougall, Myers).

Un număr de psihologi urmează direcţia deschisă de Galton în studiul diferenţelor intelectuale. Cel mai semnificativ este psihologul Alfred Binet, care inventează, cu Simon, o scală practică pentru identificarea copiilor care pot beneficia de şcolarizare, iniţiind astfel testarea inteligenţei. Binet introduce în plus faţă de demersul lui Galton ideea de a cere subiecţilor să răspundă unor sarcini complexe precum definirea cuvintelor sau punerea în aplicare a unor instrucţiuni. Psihologii de până atunci măsurau performanţe psihofizice şi senzoriale simple (Weber, Fechner), precum timpul de reacţie, considerate a fi „cărămizile"inteligenţei.

Apariţia şi dezvoltarea unor instrumente psihologice sunt legate de psihologia experimentală (măsurarea diferitelor procese psihice sau psihofizice simple), de dezvoltarea treptată a şcolii de psihologie clinică de la Paris, precum şi de începutul cercetărilor de psihologie diferenţială.

2.2. EVOLUŢIA PSIHODIAGNOZEI ÎN SECOLUL XX

Binet şi Simon încep încă din 1900 primele interviuri sistematice pentru măsurarea inteligenţei. încă dinainte de începutul secolului XX, ei scriu despre scopul psihologiei individuale ca studiu al omului, evidenţiind trei probleme importante ale acesteia:

- studiul diferenţelor individuale ale proceselor psihice fără să se studieze special relaţia cu indivizii care prezintă aceste procese;
- studiul diferenţelor individuale ale indivizilor izolaţi sau grupelor de indivizi;

- studiul raportului diferitelor procese psihice între ele în cadrul aceluiaşi individ.

Ei evocă două metode de studiu a raporturilor diferitelor procese psihice la un acelaşi individ normal:

1. să faci procesul să varieze şi să examinezi dacă această variaţie antrenează schimbări în alte procese la aceeaşi persoană;
2. să studiezi un anumit număr de procese psihice, alese în prealabil, la mai mulţi indivizi şi să examinezi dacă diferenţele individuale între procese sunt paralele şi corespund de la un individ la altul (relaţia între aceste procese fiind dedusă din această corespondenţă).

A doua metodă se va constitui ca una dintre primele finalităţi ale metodei testelor.

1904 Spearman introduce teoria bi-factorială a inteligenţei (factor general şi factori specifici) şi propune proceduri statistice.
1905 Jung prezintă *Experimentul asociativ-verbal* pentru a descoperi şi diagnostica complexele inconştiente.
1905 Nicolae Vaschide şi Henry Pieron, în Franţa, pun la punct tehnicile psihometrice ca metode ale psihologiei experimentale.
1905 Alfred Binet deschide un laborator într-o şcoală primară din Paris şi. împreună cu Simon, dezvoltă şi prezintă o scală de diagnostic a nivelului intelectual pentru copii normali şi anormali tip *screening.*
1908 Binet şi Simon prezintă a doua scală de inteligenţă. Inventează ideea de vârstă mentală. Are loc introducerea, în S.U.A., a scalei lui Binet. Pieron introduce termenul de comportament în Franţa. Lahy studiază selecţia conducătorilor de tramvai. în Belgia apare primul cabinet european de orientare profesională.
1910 Se creează Asociaţia Internaţională de Psihanaliză. J.M.Cattell prezintă în „Mind" prima serie de 10 instrumente psihologice, *teste mentale* experimentate pe studenţii Universităţii Pennsylvania.
1911 apare a III-a scală de inteligenţă Binet - Simon.
1912 William Stern inventează ideea de quotient intelectual, apelând la evaluarea nivelului actual de inteligenţă prin distanţa dintre vârsta reală şi vârsta mentală, evaluare considerată de autor predictivă pentru performanţele ulterioare ale subiectului testat.
1916 Izard şi Simon, prima scală de dezvoltare pentru prima copilărie.
1917 Utilizarea testului *Army Alpha şi Beta Examination* pentru testarea pe populaţii largi ale armatei S.U.A.. primul test colectiv nonverbal. Gesell realizează un sondaj mental al şcolilor elementare.
1918 Woodworth dezvoltă primul inventar major de personalitate, pentru *screening-ul* recruţilor din armata americană în privinţa normalităţii şi dificultăţilor de adaptare.

După primul război mondial, testele devin o practică de masă, cu cea mai largă răspândire în S.U.A. (selecţia copiilor dotaţi, selecţia şi orientarea recruţilor) şi Franţa (orientarea şcolară şi problematica vârstei mentale a copilului şi adultului).
Rorschach publică *Psihodiagnosticul,* introducând utilizarea a zece planşe cu pete de culoare pentru evaluarea bolii mentale.
1922 J.M. Cattell stabileşte Corporaţia Psihologică, o primă firmă de publicare de teste şi consultanţă.
1923 Se publică *Testul de Dezvoltare mentală Stanford,* inaugurând în America testarea elevilor de şcoală.
1926 Florence Goodenough publică testul *Desenaţi un Om,* test pentru evaluarea abilităţilor mentale.
1926 E.K.Strong publică *Testul de interese rocaţionale.*
1928 Hartshorne şi May publică prima serie de cercetări. *Studii privind înşelăciunea,* care includ teste situaţionale şi care prezintă o consistenţă scăzută a „trăsăturilor de caracter" de-a lungul diferitelor situaţii.
1934 Moreno publică *Cine va supravieţui,* introducând sociometria, măsurarea modelelor atitudinilor şi acţiunilor de grup.
1935 Thorndike dezvoltă tehnica de analiză factorială şi o utilizează pentru determinarea abilităţilor primare.
1936 Doll publică *Scala de maturitate socială Vineland* pentru evaluarea funcţionării retardului mental.
1936 în U.R.S.S. se interzice folosirea testelor psihologice.

1938L. Bender introduce *Testul Vizual - motor de gestalt Bender,* pentru evaluarea gradului de maturizare, a deteriorării creierului şi a personalităţii.
1938 H. Murray şi colaboratorii publică *Explorări ale personalităţii,* incluzând date despre tehnici de măsurare, inclusiv *Testul de apercepţie tematică,* T.A.T., început în 1935.
1939 L. Frank introduce termenul de „tehnici proiective" pentru probele cu stimuli ambigui asupra cărora subiectul „proiectează" stările şi nevoile interioare
1942 Hathaway şi McKinley publică *Inventarul multifazic Minnesota, M.M.PI.,* primul instrument complex de evaluare diferenţială a patologiei mentale.
1949 A. Davis consideră că testele de inteligenţă suportă influenţa gradului de cultură al subiectului, deci sunt incorecte faţă de cei cu nivel economic scăzut.
1954 P. Meehl publică Predicţie *clinică versus statistică,* demonstrând că formulele statistice egalizează sau îmbunătăţesc evaluările experţilor.
1955 Meehl şi Cronbach propun validitatea de construct.
1957 Edwards demonstrează influenţa dezirabilităţii sociale asupra răspunsului la itemii de personalitate, iniţiind controversa asupra set-ului de răspuns.
1957 Cronbach şi Geser lărgesc teoria testelor incluzând luarea deciziei şi problema valorilor instituţionale vs. individuale.
1957 Campbell şi Fiske propun validarea convergentă şi discriminatorie şi procedura multitrăsătură - multimetodă în dezvoltarea testelor.
1959Guilford propune o structură tridimensională a intelectului care cuprinde 120 de abilităţi speciale măsurabile prin proceduri specifice
1963Glase propune testarea cu referinţă la criteriu, opusă testării cu referinţă la normă.
1966 Asociaţia Psihologilor Americani publică Standardele pentru Testele şi *Manualele Educaţionale şi Psihologice.*
1969 A.P.A. introduce ghidurile pentru testarea celor „dezavantajaţi".
1969 Interpretarea automatizată a testelor, raportul lui Butcher asupra M.M.P.I.
1969 Utilizarea computerelor interacţionând cu clienţii, Super „Consilierea asistată de computer".
1969 În S.U.A., cazul Griggs vs. Duke Power, Curtea Federală decide că testele utilizateîn selecţia profesională trebuie să demonstreze relevanţa pentru profesia dată.

În decursul anilor sfârşitului de secol au evoluat mult atât metodologia construirii şi aplicării testelor psihologice, cât şi concepţia asupra psihodiagnozei în sensul evaluării formative.

Orientările contemporane se diferenţiază prin nivelul aparaturii statistice din ce în ce mai sofisticate care în prezent permite cuantificări nu numai ale dimensiunilor izolate, ci şi ale interacţiunilor simple sau complexe dintre acestea. Astfel, în măsurarea personalităţii s-a Irecut de la modelele uni şi bidimensionale la modele factoriale circumplexe cu trei sau mai multe dimensiuni. Dezvoltarea cercetărilor privind analiza de item, fidelitatea şi validitatea au introdus noi viziuni asupra calităţilor metrologice ale testelor psihologice şi utilizării lor în diferite tipuri de aplicaţii. S-au introdus periodizarea etalonării şi reetalonării testelor, precum şi necesitatea reviziei. S-au construit şi experimentat în funcţie de diferite tipuri de teorii asupra abilităţilor umane şi a personalităţii, o largă paletă de instrumente de testare. S-au dezvoltate standarde de testare psihologică obligatorii în diferitele state ale lumii.

2.3 CONTINUITATE Şl DISCONTINUITATE ÎN EVOLUŢIA PSIHODIAGNOSTICULUI ÎN ROMÂNIA

Nu s-a realizat încă o lucrare de istorie a psihologiei care să se refere explicit la evoluţia psihodiagnozei. Lucrări importante în tematica dezvoltării ştiinţei psihologice în România ante şi postbelică au scris M.Bejat, 1972, T.Herseni, 1980, I. Mânzat, 1994, G.Nicola, 2002. Primele cercetări în care s-au utilizat instrumente şi probe psihologice sau psihofiziologice s-au desfăşurat prin intermediul laboratoarelor de psihologie experimentală. In România, primul laborator de psihologie a fost deschis la laşi în 1893, sub conducerea lui Eduard Gruber. Doctor în
psihologie, Gruber deschide în 1893 şi primul curs de Psihologie experimentală. În ultimul deceniu al secolului XIX au apărut mai ales lucrări de cercetare în limba română sau în limbi străine, dar şi lucrări relevante pentru spiritul novator european în domeniul experimental: *Introducere în psihofizică.* Şt. Michlăiescu *Principii de psihologie,* (C.Leonardescu), *Problemele psihologiei,* (C.Rădulescu-Motru). De numele lui Rădulescu-Motru se leagă şi primul curs de Psihologie experimentală la Bucureşti - *Elemente de psihologie experimentală.* O dată cu anul 1906 obţine

fonduri pentru înfiinţarea primului laborator de psihologie experimentală la Universitatea din Bucureşti.

În România, perioada interbelică a semnificat un început promiţător. Profesorul Rădulescu-Motru reluase în 1919 Cursul de psihologie la Universitatea din Bucureşti. In psihologie oferă un model de interpretare psihologică; de exemplu, modelul de explicare a transferului de energie de la forme simple la forme complexe, cristalizare teoretică ce a avut reverberaţii semnificative în psihologia epocii.

In 1919 se înfiinţează Universitatea din Cluj, iar profesorul FI. Ştefănescu-Goangă, şef al catedrei de psihologie, edifică Institutul de cercetări în 1921 şi începe demersul pentru înfiinţarea unei secţii de psihologie aplicată în cadrul căreia psihodiagnoza apare în prim plan (metode, teste experimentate şi etalonate etc.). De asemenea, iniţiază o serie 'de servicii de psihologie aplicată la calea ferată, armată, poştă, instituţiile juridice şi de reeducare, în sănătate şi educaţia specială, dar şi în educaţia diferenţiată orientată spre performanţă. In 1929, 1933 publică S*elecţiunea capacităţilor şi orientarea profesională*. Formează şi susţine dezvoltarea ca cercetători şi practicieni a unor discipoli importanţi precum: N. Mărgineanu, D. Tudoran, Al. Roşca, M. Beniuc. în problema specifică a psihodiagnozei, centrarea cercetărilor clujene este pe etalonarea şi adaptarea testelor psihologice, muncă vastă în care conlucrează şi profesori şi psihologi din Bucureşti: Gh. Zapan, G. Bontilă şi I. Nestor.

Perioada de după cel de al II-lea război mondial a semnificat stagnare şi regres. Nu s-au realizat revizii ale metodologiei psihodiagnozei în ultimii 50 de ani. Sunt relativ puţine încercările de a crea noi teste. Au apărut în perioada postbelică puţine lucrări privind testarea psihologică, cele publicate având fie un caracter didactic, fie un caracter general. Au existat preocupări în cadrul laboratoarelor industriale de realizare a unor monografii profesionale şi a unor psiho-profesiograme aplicabile în selecţia şi orientarea profesională. Laboratoare au existat în industrie, aviaţie, armată, clinică, precum şi în centrele universitare din Bucureşti, Cluj, Iaşi şi Timişoara. Consilierea şcolară şi profesională, începută strălucit în perioada interbelică, s-a diminuat treptat până la dispariţie. După 1983, o dată cu conul de umbră în care au fost aruncate psihologia şi munca psihologilor din România, activitatea aplicativă şi formativă s-a restrâns deosebit de mult.

În prezent este marcantă tendinţa de a computeriza activitatea de evaluare psihologică (aplicare, scorare, interpretare).

Se răspunde mai puţin semnificativ tendinţei actuale în plan internaţional de a se reduce perioada dintre ultima revizie a testului şi cea curentă.

S-a început construirea unor noi instrumente. Dacă. în plan internaţional, testele de psiholingvistică Big Five, precum şi Bateriile multidimensionale de aptitudini sunt o preocupare actuală, în România au început demersuri paralele (testul ABCD-M de tip Big Five), dar se accentuează mai ales pe teste specifice şi situaţionale de selecţie profesională adecvate unor locuri de muncă particulare (aviaţie etc). S-au realizat mai ales chestionare cu o aplicabilitate restrânsă.

In plan internaţional, se construiesc noi instrumente de testare pentru arii aplicative precum diagnoza formativă, psihologia clinică, psihologia comportamentului sau psihologia ecologică; în România această preocupare este relativ restrânsă şi depinde mai ales de interesul specific al unui profesionist.

în plan internaţional, tendinţa este de a completa echipamentul lestului cu noi studii privind faţetele validităţii şi utilitatea acestuia; manualul tehnic include aceste informaţii privind caracteristicile metrologice, dar şi rezultate ale cercetărilor privind aplicaţiile şi interpretarea scorurilor. Apar ghiduri privind utilitatea aplicării testelor în diferite arii aplicative. Aceste aspecte sunt neglijate şi nu există încă preocupări metodologice convergente şi nici pentru publicarea manualului testului cu cercetările aferente. Apar sporadic articole în revistele de specialitate privind unele experimentări de teste (E.P.I., C.P.I.. etc).

Există slabe preocupări în România privind creşterea valorii fidelităţii şi a validităţii probelor utilizate (N. Mitrofan, 2001). De asemenea, slabe preocupări de experimentare şi standardizare a testelor pe întreaga populaţie (excepţie grupul de cercetare Cluj pentru testele C.P.I. Wonderlic, E.P.Q., şi grupul de cercetare Bucureşti, testul *screening* Denver, NEO-PI-R, M.B.T.I.).

Există slabe preocupări pentru luarea măsurilor de protecţie legală a instrumentului psihodiagnostic şi prevenirea situaţiilor în care să se limiteze/permită accesul la teste profesionale şi la interpretarea rezultatelor, pentru persoane neformate profesional. În acest sens, o importanţă decisivă o are în prezent preocuparea Asociaţiei psihologilor din România de a impune nu numai recunoaşterea profesiei, ci şi. mai ales, a unei Legi privind exercitarea profesiunii de psiholog, constituirea unui

Consiliu al psihologilor care să impună respectarea metodologiilor în evaluarea psihologică şi aplicarea Codului deontologic al profesiunii.

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition),*Routledge, London and New York,
Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului*, Ed. Fundatiei Romania de Maine, Bucuresti,
Mitrofan Niculae,2001- Psihometria şi direcţiile ei de dezvoltare la început de mileniu, în Zlate, M. (coord.), *Psihologia la răspântia mileniilor*. Iaşi: Polirom.
Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare*, Editura Polirom, Iaşi
Urbina Susana, *2004- Essential of psychological testing,* John Wiley & Sons, Inc. Hoboken, New Jersey,
COLEGIUL PSIHOLOGILOR DIN ROMÂNIA (2005 şi 2006), *Acte normative*.

3. CODUL DEONTOLOGIC AL PRACTICĂRII PSIHODIAGNOZEI

3.4. Legi, standarde şi reglementări privind testarea educaţională şi psihologică. Situaţia din ţară noastră.
3.5. Responsabilităţi în construirea testelor, publicare şi comercializarea acestora
3.6. **Responsabilităţi în aplicarea testelor şi utilizarea rezultatelor**

1.1. LEGI STANDARDE ŞI REGLEMENTĂRI PRIVIND TESTAREA EDUCAŢIONALĂ ŞI PSIHOLOGICĂ

Asociaţiile psihologilor din diferite tări au construit o serie de norme asamblate în coduri deontologice valabile pentru conduita profesională a psihologului şi a oricarei persoane implicate în funcţii profesionale de psihologie.
Comportamentul profesionist şi responsabilităţile în utilizarea modalităţilor de evaluare psihologică pot fi privite din trei perspective:

- Probleme legale
- Probleme etice
- Probleme profesionale

În consecinţă discutăm de:

- legi,
- standarde etice şi
- standarde profesionale.

Acestea formează un continuum de standarde ce privesc diferite laturi ale comportamentului profesional. Documentele legale despre testare sunt în general mai vagi dar influenţează standardele profesionale privind conduita şi practicile profesionale în evaluare şi testare. Codurile etice şi ghidurile profesionale, principiile şi standardele aferente sunt dezvoltate pentru a educa şi ghida profesioniştii în activităţile concrete.

Primul cod etic
Primul cod etic formal pentru evaluarea psihologică a fost adoptat de Asociaţia Psihologilor Americani în anul 1952. Cu titlu exemplificativ, din cele 100 de principii etice, 18 se adresează utilizării testelor psihologice, acoperind următoarele domenii:

- Calificarea celor ce utilizează teste- 3 principii
- Responsabilităţile psihologului- 4 principii
- Responsabilităţile şi calificările celui care publică teste- 3 principii
- Condiţia ca testul să fie considerat terminat- 1 principiu
- Modul de descriere a testului în manual şi publicaţii- 5 principii
- Securitatea materialelor testului- 2 principii

În anii 1980 majoritatea asociaţiilor psihologice din lume au adoptat standarde etice şi coduri profesionale care se referă şi la domeniul măsurării şi evaluării psihologice.

1.1.1. TIPURI DE STANDARDE

Informaţiile privind istoricul reglementărilor în alte tări sunt utile pentru a evalua unde ne aflăm acum şi cum trebuie să privim activitatea psihologică în context social, învăţând din experienţele relevante pe plan mondial. Tradiţia cea mai mare în domeniu o are activitatea psihologică din S.U.A.
În anul 1985, în S.U.A. au fost elaborate şi adoptate un set de norme etice şi standarde de utilizare a testelor psihologice şi de practicare a examinărilor psihologice sub titulatura: *„Standards for Educational and Psychological Testing”* . La elaborarea lor au contribuit trei instituţii The American Educational Research Association(A.E.R.A.), The American Psychological Association (A.P.A.) şi The Nacional Council on Measurement in Education (N.C.M.E.). Standardele publicate pentru prima oară în 1952 au fost revizuite succesiv în 1966, 1974, 1985 şi 1999. În aceste standarde sunt vizate atât caracteristicile psihometrice ale testelor cât şi comportamentul profesional al psihologului, insistându-

se asupra faptului că acestora le revine responsabilitatea profesională de a apăra bunăstarea psihologică a persoanelor testate.
Legile care se referă la testare au ca principal scop protejarea anumitor segmente ale publicului de abuzurile specifice.
Standardele etice şi codurile sunt mai specifice şi încearcă să stabilească norme mai înalte pentru o gamă mai largă de profesionişti, care trebuie să se raporteze în activitatea lor în primul rând la codul etic.
Standardele profesionale, principiile şi ghidurile sunt orientate în general spre probleme mai tehnice pentru a-i ghida pe cei care utilizează testul faţă de aplicaţii specifice şi utilizări ale evaluărilor. Spre exemplu aceste standarde răspund unor problematici legate de dezvoltarea, validarea şi utilizarea evaluărilor, de utilitatea testului, de corectitudinea sa, de felul cum face faţă constrângerilor practice.

În completarea Standardelor din 1985 au apărut prevederi specifice pentru domenii particulare, spre exemplu:
Special Guidelines(APA, 1981)- pentru psihologii clinicieni, organizaţii industriale, şcoli sau în consiliere
Principles for The Validation And Use of Personnel Selection Procedures (Society For Industrial And Organizational Psychology, 1987)- pentru selecţia profesională
Guidelines For Computer-Based Tests And Interpretations (APA; 1986)- pentru folosirea calculatorului în testarea psihologică
Ethical Principles and Code of Conduct(APA, 1992)- cod deontologic specific evaluarii psihologice.
Câteva din aspectele cuprinse în codul deontologic(APA, 1992):

- Evaluarea psihologică şi psihodiagnosticul trebuie să se facă numai într-un cadru profesional, de către utilizatori competenţi şi cu experienţă în domeniu, cu ajutorul unor teste potrivite
- Pentru proiectarea şi selectarea tehnicilor adecvate se vor folosi proceduri ştiinţifice
- Interpretarea rezultatelor şi explicarea acestora se va face cu mare atenţie
- Se va acţiona pentru păstrarea securităţii testelor

Prezentăm spre exemplificare problematica abordată de Standardele A.P.A. privind calificarea utilizatorilor de teste(adoptate în anul 2000), care vin în completarea şi specificarea celorlalte standarde, şi a căror necesitate a apărut din observaţia că majoritatea dificultăţilor legate de teste provin din insuficienta calificare a utilizatorilor. Câteva probleme din cuprinsul acestora:
Cunoştinţe şi abilităţi esenţiale pentru utilizatorii de teste

- Cunoştinţe privind psihometria şi măsurarea psihologică
- Selectarea testelor, procedurilor adecvate
- Proceduri de administrare a testelor
- Influenţa variabilelor etnice, rasiale, culturale, de sex, de vârstă şi lingvistice
- Testarea persoanelor cu disabilităţi
- Experienţa supervizată

Calificări necesare pentru contexte specifice

- În contextul ocupaţional
- În contextul educaţional
- În contextul consilierii ocupaţionale, vocaţionale
- În contextul sănătăţii
- În contextul justiţiei

În ceea ce priveşte capitolul "Cunoştinţe privind psihometria şi măsurarea psihologică", standardele precizează ariile de cunoştinţe, din care redăm o parte:

- Cunoştinţe despre teoria clasică psihometrică şi teoria răspunsului la item
- Statistică descriptivă- distribuţii de frecveţe, măsuri ale tendinţei centrale, măsuri ale variaţiei, indicele de relaţie(corelaţia datelor)
- Scale, scoruri şi transformări- tipuri de scale(nominale, ordinale, de interval, de proporţie), tipuri de scoruri(brut, transformat- percentile, standard, normalizate)

- Fidelitatea şi erorile de măsură- surse de variabilitate sau erorile de masurătoare(caracteristicile examinatului, caracteristicile testului, caracteristicile constructului şi intenţia de utilizare a scorurilor, caracteristicile şi comportamentul examinatorului, caracteristicile mediului în care se face testarea, procedurile de administrare a testului, acurateţea scorării), tipuri de fidelitate şi adecvarea lor pentru diferite tipuri de teste şi utilizări ale testelor(fidelitatea test-retest, fideliatea calculată prin formele paralele sau alternative ale testelor, consistenţa internă, fidelitatea interevaluatori), eroarea standard a măsurătorii(SEM)
- Validitatea şi semnificaţia scorurilor la teste- tipuri de dovezi care contribuie la validitatea de construct (de conţinut, de criteriu, convergentă, discriminativă), interpretarea normativă a scorurilor la teste(tipuri de norme şi relevanţa lor pentru interpretarea rezultatelor – scoruri standard sau scoruri percentile), caracteristici ale grupului normativ şi limite în generalizarea rezultatelor, tipuri de referenţi ai scorurilor (norma, domeniul- criteriul, autoreferenţiale- scalele ipsative), tabelele de expectanţe.

SITUAŢIA DIN ŢARA NOASTRĂ

În România în anul 2004 a fost adoptată şi a intrat în funcţiune legea care reglementează profesia de psiholog, cu denumirea completă: *„Legea nr. 213 din 27 mai 2004, privind exercitarea profesiei de psiholog cu drept de liberă practică, înfiinţarea, organizarea şi funcţionarea Colegiului Psihologilor din România”*.

Această lege trebuie să fie cunoscută de orice psiholog din România, incluzându-i şi pe cei care se află în proces de formare (studenţii, masteranzii). În baza ei s-a înfiinţat Colegiul Psihologilor din România, iar în cadrul Colegiului funcţionează patru comisii aplicative, şi încă două comisii, una metodologică şi una deontologică, respectiv:

- Comisia de psihologie clinică, consiliere psihologică şi psihoterapie;
- Comisia de psihologie a muncii, transporturilor şi serviciilor;
- Comisia de psihologie educaţională, consiliere şcolară şi vocaţională;
- Comisia de psihologie pentru apărare, ordine publică şi siguranţă naţională.
- Comisia metodologică
- Comisia de deontologie şi disciplină.

Comisia metodologică are atributiuni mai apropiate de activitatea psihodiagnostică este elaborând deja „Normele de avizare a metodelor şi tehnicilor de evaluare şi asistenţă psihologică”. Conform art.3 (1) Comisia Metodologică a Colegiului Psihologilor din România, prin normele de avizare prezente, stabileşte standardele de calitate şi procedurile de avizare pentru metodele şi tehnicile de evaluare şi asistenţă psihologică. Pe de altă parte, conform art. 3(2) „Comisiile aplicative din cadrul Colegiului Psihologilor din România, prin normele de avizare specifice, atestă competenţele profesionale ale psihologilor de a utiliza metode şi tehnici de evaluare şi asistenţă psihologică. Toate comisiile de specialitate vor avea în vedere competenţele generale de utilizare a testelor psihologice precizate în documentul de faţă”.

De o mare importanţă sunt precizările referitoare la dreptul de utilizare a testelor psihologice. Cei care folosesc testele trebuie să fie în posesia unor competenţe bine precizate, identice cu cele prezentate în Standardele APA pentru calificarea utilizatorilor de teste(2000), care se obţin prin procesul de formare profesională, însemnând şi anii de studiu şi de practică din facultate, dar şi procesul de formare postuniversitară.

Categoriile majore de competenţe vizează următoarele aspecte etice în aplicarea testelor:

- adoptarea unei conduite profesioniste, în acord cu normele deontologice ale psihologului şi cu respectarea legilor internaţionale şi naţionale privind copyright-ul (legea drepturilor de autor şi a drepturilor conexe – nr. 8/1996);
- utilizarea doar a acelor teste pentru care au competenţa necesară;
- asumarea responsabilităţii pentru modul de utilizare a testelor;

- asigurarea securităţii pentru testele utilizate, astfel încât ele să nu-şi piardă calităţile din cauza deconspirării publice a conţinutului ori a mecanismelor de cotare;
- asigurarea confidenţialităţii rezultatelor;
- acordul scris sau în formă electronică de includere a rezultatelor în baza de date a utilizatorului probei psihologice.

De asemenea, această comisie a elaborat o serie de norme referitoare şi la modul de comercializare a testelor.

1.2. RESPONSABILITĂŢI PRIVIND CONSTRUIREA TESTELOR, PUBLICAREA ŞI ACCESUL LA ACESTEA

După standardele din ghidurile tehnice testul psihologic este:

- Un instrument de interpretare utilizat de psihologi, formând o bază pentru o inferenţă care depăşeşte conţinutul itemului
- Se bazează pe principii ştiinţifice
- Are proprietăţi psihometrice cunoscute şi evaluate
- Poate fi folosit pentru o varietate de întrebuinţări
- Poate fi folosit de profesionişti care nu sunt psihologi şi chiar utilizat de profesionişti care nu sunt psihologi, dar care trebuie să fie capabili să-l folosească pentru scopuri psihologice, respectând principiile etice ale practicii psihologice

Standardele tehnice cuprind norme pentru construirea testului ce se referă la dovezi ştiinţifice privind validitatea, gradul de încredere şi erorile măsurătorii, cercetările privind dezvoltarea şi revizia testului, scalarea şi stabilirea normelor, comparabilitatea scorurilor, standardele de publicare.
Există două tipuri acceptate de utilizatori:

- Utilizatorul primar de teste
- Utilizatorul secundar

Utilizatorul primar este profesionistul psiholog. Utilizatorul secundar poate fi o altă persoană care are acces la teste, care nu are formare în măsurare şi evaluare, dar care, mai ales în domeniul educaţional, poate avea nevoie de teste, distorsionând sau interpretând greşit rezultatele testelor. Astfel utilizatorul primar este implicat, de regulă, în îmbunătăţirea utilizării testelor, informarea şi influenţarea şi grupurilor de utilizatori secundari. Codul pentru practici corecte în educaţie, în S.U.A. a apărut într-un astfel de context.

În privinţa calităţii instrumentelor şi a procedurilor de măsurare, normele prevăd obligaţia psihologului de a citi manualul testului pentru a determina dacă:

- Testul este adecvat scopului urmărit
- Testul este potrivit caracteristicilor celor examinaţi
- Calităţile testului sunt acceptabile
- Dacă are validitate de aspect

Privind distribuirea testelor, firmele care produc şi distribuie teste psihologice sunt obligate să acorde atenţie prevederilor legale privind controlul distibuirii testelor, numai către utilizatorii calificaţi profesional, cu diplomă de licenţă în psihologie(ex. Franţa, România), sau şi celor în curs de formare, dacă există un supervizor care îşi asumă răspunderea pentru utilizarea testelor de către cumpărător(Ex. Olanda). Asociaţia Psihologilor Americani a manifestat întotdeauna grijă pentru ca testele să fie aplicate la un nivel înalt ştiinţific, dar numai de către cei calificaţi pentru aceasta. Astfel APA clasează instrumentele psihologice în trei clase de accesibilitate profesionala : A, B şi C, clasificare care implică şi distribuirea testelor. În principiu, modelul ABC al APA stipuleaza că instrumentele psihometrice nu sunt accesibile decât utilizatorilor calificaţi.

Un test de nivel A poate fi achiziţionat de orice entitate, organizaţie sau instituţie, precum şi de orice persoană , indiferent de calificarea ei profesională . De obicei testele

de calificare A, care nu cer o anumită formare profesională , sunt destinate uzului în business sau în zona educaţională . Testele de clasa A nu sunt de obicei instrumente foarte puternice ori foarte documentate. Acestea pot fi administrate, scorate şi interpretate adecvat cu ajutorul manualului şi al unei orientări generale (de exemplu, teste de achiziţie).

Instrumentele de nivel B sunt destinate acelor indivizi care au absolvit studii universitare cu profil psihologic ori asimilate acestora(asistenţă socială, psihosociologie, psihopedagogie specială). Aceasta este o concesie făcută situaţiei din România, deoarece standardul APA stabileşte accesibilitatea instrumentelor de tip B exclusiv după absolvirea unui curs de psihometrie la nivel de Master, precum şi (condiţii cumulate!) training specializat sub supervizarea unui psiholog calificat. Testele de nivel B reclamă anumite cunoştinţe tehnice privind construirea şi utilizarea testelor, precum şi cunoştinţe din alte domenii ale psihologiei, precum: psihologia diferenţială, statistica psihologică, psihologia personalului, psihologia adaptării, orientarea vocaţională (de exemplu, teste de aptitudini, inventare de adaptare aplicabile populaţiei normale). Majoritatea inventarelor de personalitate sunt clasate B.

Instrumentele de clasa C sunt destinate exclusiv indivizilor care au obţinut titlul doctoral în psihologie sau educaţie, sau care sunt membri ai unor organizaţii profesionale care îi califică pentru uzul respectivului test, ori care lucrează sub supervizarea directă a unui psiholog calificat pentru uzul respectivului test. Aceste teste solicită cunoştinţe de înalt nivel privind testarea şi din alte domenii de suport, la care trebuie să se adauge supervizarea din partea unor specialişti în aceste domenii (de exemplu, teste proiective, teste de inteligenţă individuale). Testele clinice (MMPI şi Rorschach de exemplu) precum şi o parte din testele de abilităţi cognitive(Weschler) sunt clasate C, primele datorită experienţei profesionale necesare, cele din urmă mai degrabă datorită fineţii deontologice necesare în înţelegerea şi comunicarea rezultatelor.

Este important de remarcat că există o multitudine de factori care sunt luaţi în considerare în judecarea nivelului de calificare al unei persoane, pe lângă formarea sa profesională. Se consideră calitatea de membru în organizaţii profesionale cu activitate în zona psihologiei sau consilierii, calitatea de angajat al unei instituţii active în aceste zone, experienţa generală în testare şi în psihometrie, precum şi utilizarea intenţionată a testului.

În S.U.A., în *Ethical Principles of Psychologists* (1992) sunt făcute precizări şi mai clare privind calificările utilizatorilor de teste. Astfel, *Principiul A* afirmă că psihologii „oferă numai acele servicii şi utilizează numai acele tehnici pentru care sunt calificaţi prin educaţie, formare sau experienţă. *Standardul etic 2.02* adaugă: „psihologii se abţin de la utilizarea necorespunzătoare a tehnicilor de măsurare şi, totodată, iau măsuri pentru a preveni ca alţii să utilizeze defectuos informaţia pe care o oferă aceste tehnici”.

Cu toate măsurile luate s-a constatat că, în practică, se întâlnesc frecvente abateri. De aceea, APA, împreună cu The American Research Association, The National Council on Measurement in Education şi cu foarte multe edituri ce publicau teste, a înfiinţat The Joint Committee on Testing Practice (JCTP). Acesta a adoptat, în 1988, *The Code of Fair Testing Practices in Education*, care descrie obligaţiile celor care construiesc teste, precum şi ale celor ce le aplică, în patru arii: a) dezvoltarea/selectarea testelor; b) interpretarea scorurilor; c) strădania pentru corectitudine; d) informarea subiecţilor.

1.3. RESPONSABILITĂŢI ÎN APLICAREA TESTELOR ŞI ÎN UTILIZAREA REZULTATELOR

Pentru activitatea specifică testării psihologice (aplicarea, administrarea) normele naţionale şi internaţionale reglementează:

- Calificarea examinatorului
- Calitatea procedurilor şi instrumentelor de măsurare
- Obţinerea consimţământului subiectului pentru examinarea psihologică
- Confidenţialitatea rezultatelor

- Modul de comunicare a rezultatelor
- Specificarea efectului discriminatoriu al testului asupra grupelor minoritare

O să luăm în discuţie pe cele mai importante dintre acestea.
Calificarea examinatorului este specificată, spre exemplu în S.U.A., în Standardele pentru calificarea examinatorilor, aprobate şi publicate în anul 2000, înfăţişate pe scurt în capitolul anterior. În plus trebuie menţionat că fiecare catalog care prezintă teste specifică şi nivelul de competenţă necesar cumpărării şi utilizării fiecărui test. Utilizatorii trebuie să aplice testele numai conform indicaţiilor date de autorii lor. Despre calitatea procedurilor de examinare se specifică interdicţia de a examina în condiţii necorespunzătoare şi în situaţia în care subiecţii prezenţi la examinare nu sunt pregătiţi sau nu-şi dau consimţământul.
Obţinerea consimţământului include obligaţia psihologului de a explica motivul examenului, în unele ţări acordul fiind realizat într-o formă scrisă, semnată de ambele părţi. De aceea se face adăugirea- consimţământ informat, însă, pentru a nu afecta rezultatele examinării, nu se dau explicaţii detaliate despre test, când intruziunea în viaţa particulară este minoră.
Confidenţialitatea rezultatelor se referă la accesul strict limitat pentru informaţia furnizată de testul psihologic. În afară de câteva excepţii- când persoana este un pericol pentru sine şi pentru alţii, sau este vorba de selecţie psihologică- informaţia nu va fi comunicată altei persoane. Chiar şi atunci când trebuie comunicată- acces au managerul, sub forma interpretată, medicul, avocatul, psihoterapeutul care se ocupă de persoană. Datorită surselor de eroare ce pot afecta semnificaţia rezultatelor subiectului i se poate da posibilitatea de a corecta ceea ce consideră a fi o imagine nefavorabilă sau discordantă cu performanţele sale reale. Trebuie evitată afişarea rezultatelor examenului psihologic, recomandându-se afişarea doar a persoanelor admise.
Comunicarea rezultatelor se face doar sub forma lor prelucrată şi interpretată, sub formă de raport psihologic, conform cerinţelor beneficiarului. Raportul trebuie să fie clar, fără ambiguităţi, evitându-se utilizarea termenilor tehnici, care pot antrena erori de înţelegere din parte unor neprofesionişti. În raport se explică semnificaţia rezultatelor şi posibilele implicaţii, în termeni comportamentali, relevanţi pentru scopul examenului, cu posibilitatea unei predicţii pe un termen scurt, în general şase luni, pentru anumite contexte. În general se evită formularea în termeni de certitudine, preferându-se formulări relativizante, de genul: „în general”, „în anumite contexte”, „este posibil ca”, „este probabil să”. Datele comunicate trebuie să reflecte situaţia cât mai obiectiv, fără păreri subiective sau strict personale, ţinând cont de toate informaţiile oferite de test, situaţia de testare, observaţia pe timpul testării şi caracteristicile psihometrice ale testului.
Posibilitatea discriminărilor intervine mai ales când proba este utilizată în selecţia profesională şi se utilizează un singur prag de admisibilitate. Dacă există grupuri minoritate care obţin scoruri mai mici decât alte grupuri, persoanele care fac parte din acestea au şanse mai mici să fie selectate şi astfel pot fi discriminate. În acest sens sunt prevăzute şi standarde pentru aplicaţiile particulare ale testelor, spre exemplu testarea lingvistică a minorităţilor etnice şi testarea persoanelor cu disabilităţi, dar nu numai. Însă situaţia discriminării reale apare numai dacă persoanele ar obţine şi performanţe echivalente în munca respectivă, spre exemplu în selecţia de personal, iar rezultatele la teste ar fi influenţate doar de un criteriu care nu are relevanţă pentru performanţa viitoare (spre exemplu diferenţele datorate sexului la un examen psihologic la aviaţie, sau pentru postul de secretară când la examen se prezintă un bărbat, sau cele legate de limba de examinare- atunci când performanţa nu depinde de abilităţile lingvistice sau de cultură). Pentru a preîntâmpina astfel de erori se calculează ecuaţia de regresie pentru grupul majoritar şi pentru cel minoritar, iar dacă este aceeaşi atunci testul nu face discriminări, deci rezultatele la test prognosticheză performanţe mai slabe. Dacă diferă atunci se pot fixa praguri de admisibilitate mai scăzute pentru grupurile minoritate respective (cutoff score).

În ceea ce priveşte testele educaţionale, preocupări destul de recente includ în standardele de utilizare reglementări privind:

- Notificarea anticipată a standardelor de conţinut dar şi de nivel al performanţei la testele respective
- Determinarea posibilităţii ca elevii să-şi însuşească cunoştinţele şi abilităţile cerute

- Antrenarea în scopul atingerii succesului- familiarizarea cu tipul de sarcini, inclusiv cu modalitatea practica de evaluare(de exemplu familiarizarea cu calculatorul, tezele unice care anticpează examenele)
- Evaluarea reflectă practicile curente de instruire şi curriculare
- Existenţa unei încrederi supra sau sub dimensionate în scoruri- informaţiile trebuie suplimentate cu alte informaţii din alte surse, iar performanţa trebuie interpretată într-un context mai larg
- Determinarea şi eliminarea posibilităţilor de a „învăţa” testul sau de a trişa
- Luarea în considerare a fundalului cultural şi social al celui testat
- Excluderea unei largi proporţii de elevi sau studenţi care prezintă dificultăţi sau disabilităţi lingvistice din studiile şi normele care monitorizează progeresul educaţional şi permit comparaţii sistematice(testările pe scară largă), duce la imagini şi standarde deformate

Toate aceste aspecte privind codul deontologic al psihodiagnosticianului intervin în percepţia socială şi creează imaginea unei ştiinţe eficiente şi obiective sau a ineficienţiei şi amatorismului, promovând astfel încrederea şi utilizarea la nivel social sau dimpotrivă neîncrederea şi minimalizarea socială. De aceea este important să fie cunoscute şi aplicate de către toţi cei care profesează în domeniu.

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition),*Routledge, London and New York,
Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului,* Ed. Fundatiei Romania de Maine, Bucuresti,
Mitrofan Niculae- Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare*, Editura Polirom, Iaşi
Urbina Susana, *2004- Essential of psychological testing,* John Wiley & Sons, Inc. Hoboken, New Jersey,
COLEGIUL PSIHOLOGILOR DIN ROMÂNIA (2005 şi 2006), *Acte normative*.

4. TESTUL PSIHOLOGIC INSTUMENT DE BAZA ÎN EVALUAREA PSIHOLOGICĂ

4.5. Definiţii ale testului psihologic din diferite perspective
4.6. Clasificări ale testelor psihologice
4.7. Calităţile generale ale testelor şi echipamentul testului psihologic
4.8. Selectarea celor mai adecvate teste pentru o situaţie

DEFINIŢII ALE TESTULUI PSIHOLOGIC DIN DIVERSE PERSPECTIVE

Etimologie
Cuvântul test provine din latinescul **testum**, care semnifică o oală de pământ, capacul acesteia. Cuvântul test este legat de cuvântul cap, testa. Ca şi conotaţie a insemnat- încercare, probă, examen critic, piatră de încercare, mărturie.(Minulescu, M., 2003).
Termenul de **test** este legat de cel de diagnostic psihic, care, la rândul lui, provine de la cuvântul grecesc **diagnosticos**, însemnând „apt de a recunoaşte". Până la un anumit punct, diagnosticul psihic este similar cu diagnosticul medical, deoarece există o multitudine de caracteristici specifice.
Observaţie: distincţia făcută de unii autori între noţiunea de test şi chestionar de personalitate
Accepţiunea încetăţenită- test

Vom trece în revistă mai multe definiţii ale testului psihologic, oferite de autori străini şi români:

1.Conform definiţiei de lucru din *„Standardele APA privind calificarea utilizatorilor de teste(2000)",*"testul este o **procedură de măsurare** pentru evaluarea **caracteristicilor psihologice** ale unei persoane, în care un **eşantion** din comportamentul examinatului este obţinut, evaluat, scorat, prin utilizarea unui **proces standardizat**".

2. Anne Anastasi şi Susana Urbina(1997) au o definiţie foarte asemănătoare: „Un test psihologic este în mod esenţial, o măsură **obiectivă** şi **standardizată** a unui **eşantion de comportament."**

3. Susana Urbina(2004) reia definiţia şi o completează " Un test psihologic este o **procedură sistematică** utilizată pentru a obţine **eşantioane de comportament**, **relevante pentru funcţionarea cognitivă sau afectivă** a persoanei, şi pentru **a transforma în scoruri şi a evalua** aceste eşantioane în concordanţă cu anumite **standarde**.

Merită subliniate şi definiţii ale altor autori importanţi în câmpul psihodiagnosticului.

3. L.G.Cronbach (1966)(apud Minulescu, M., 2003): „Testul este o procedură sistematică de a compara comportamentul a două sau mai multe persoane"

4. Pierre Pichot(1967): „Testele sunt instrumente de lucru **standardizate** servind de **stimuli pentru un comportament**, care, la rândul lui, poate fi evaluat prin **compararea statistică** cu comportamentul altor subiecţi aflaţi în aceeaşi situaţie, ceea ce permite **clasarea** subiectului examinat, fie cantitativ, fie tipologic".

5. Paul Oléron: „Testul este un instrument fundamental al psihologiei aplicate ce se caracterizează prin posibilitatea de **a examina numeroşi indivizi în situaţii uniformizate**".

6. A. Rey: „Testele psihologice sunt **procedee standardizate**, construite astfel încât să provoace la subiecţii investigaţi **reacţii înregistrabile** iar materialul obţinut prin intermediul acestor teste se estimează prin referinţă la **valorile etalon**"

7. M. Roşca (1972): „Testul este **o probă sau o serie de probe**, construite în scopul stabilirii **prezenţei (sau absenţei) unui aspect psihic**, a particularităţilor de manifestare comportamentală sau a gradului de dzvoltare psihică".

8. Grand dictionnaire de la psychologie (1994): „Testul este o **probă,** utilizată cu precădere în psihologia diferenţială, care **permite descrierea comportamentului** unui subiect într-o **situaţie definită precis**, prin **raportare la comportamentul unui grup precis** de subiecţi, plasaţi în aceeaşi situaţie standard. Descrierile se fac de obicei sub formă numerică".

8. Tot A. Anastasi (1992): „Testul este o măsură **obiectivă** şi **standardizată** a unui **eşantion de comportament”**.

9. Klausnitzer(1992): „Testul psihologic este o situaţie s**tandardizată** care generează un **comportament semnificativ**.”

Conceptele cheie în aceste definiţii se referă, majoritatea, la obiectivitate, standardizare şi măsurarea unui eşantion de comportament. Vom proceda la o detaliere a acestor aspecte.

Măsură obiectivă înseamnă, în principal, două aspecte: a) administrarea, cotarea şi interpretarea scorurilor sunt obiective în măsura în care acestea sunt independente în raport cu judecăţile subiective ale examinatorului: b) obiectivitatea testului reiese din faptul că determinarea nivelului de dificultate al unui item ori a întregului test se bazează pe proceduri empirice, obiective şi nu pe nivelul de apreciere (deci, subiectivitate) a celui ce construieşte un test.

Al doilea concept, **măsură standardizată,** înseamnă, de asemenea**,** mai multe aspecte: a) standardizarea implică uniformitatea procedurii în administrarea, cotarea şi evaluarea rezultatelor. Pentru ca scorurile diferitelor persoane să fie comparabile, condiţiile de testare trebuie, în mod evident, să fie aceleaşi pentru toţi. În vederea sigurării acestei uniformităţi a condiţiilor testării, constructorul testului are obligaţia să ofere informaţii şi direcţii detaliate pentru administrarea fiecărui test nou. De altfel, formularea acestor direcţii este o parte majoră a standardizării unui test nou, deoarece ele se referă la: - materialele folosite; - instrucţiunile orale (instructajul); - demonstraţiile preliminare; - modalităţile de răspuns la întrebările şi neclarităţile ridicate de către subiecţi; - orice alte detalii privind situaţia de testare; b) un alt pas important în standardizare este stabilirea *normelor*, a *etalonului*, la care raportăm, după aplicarea testului, rezultatele obţinute de către subiect. Ceea ce obţine concret la un test un subiect constituie *scorul brut* (raw score) şi el poate să exprime numărul corect de itemi, timpul cerut pentru îndeplinirea unei sarcini, numărul erorilor sau alte modalităţi de măsurare potrivite conţinutului testului. Acest scor nu spune nimic până nu îl raportăm la norme sau la etalon. (Cuvântul *norme* este preluat din limba engleză şi el este echivalent cuvântului *etalon* din limba română).

În sfârşit, conceptul **eşantion de comportament** se referă la faptul că, în acţiunea de testare, vizăm o anumită secvenţă a comportamentului sau un anumit tip de comportament, cum ar fi, de exemplu: comportament inteligent, comportament creativ, comportament motric, comportament verbal

10. O altă definiţie dată testului psihologic de către Robert Gregory (apud Mitrofan, N., 2007), aduce elemente în plus: „Testul psihologic este o procedură standardizată folosită pentru **eşantionarea comportamentului şi descrierea lui cu ajutorul categoriilor sau scorurilor**”. Plecând de la această definiţie, putem deduce faptul că majoritatea testelor prezintă următoarele caracteristici definitorii: a) procedura de standardizare; b) eşantionul de comportament; c) scoruri sau categorii; d) norme sau standarde; e) predicţia comportamentului netestat, reprezentat de eşantionul testat. *Procedura de standardizare* este trăsătura esenţială a testului psihologic. Niciun examinator nu are voie să se abată de la setul de instrucţiuni şi de precizările privind modul de aplicare a testului existente în cadrul manualului. Acestea reprezintă o garanţie privind faptul că procedurile de administrare se aplică identic de către toţi. *Eşantionul de comportament* este de interes în măsura în care el permite examinatorului să facă inferenţe asupra domeniului total al comportamentului. Un bun test trebuie să aibă capacitatea de a-i permite examinatorului să prognozeze alte comportamente şi nu pe cele reflectate de itemi. Aplicarea testelor se finalizează cu *derivarea de scoruri* sau de *categorii*. Deoarece în abordarea nomotetică se presupune că toţi oamenii posedă trăsătura sau caracteristica ce urmează a fi măsurată într-o anumită „cantitate”, scopul testării psihologice este de a estima tocmai această cantitate, exprimată numeric. De asemenea, un test psihologic trebuie să posede *norme* sau *standarde*. Rezultatele obţinute de grupul de normare sau eşantionul de standardizare sunt grupate pe mai multe clase, stabilindu-se performanţa medie şi, totodată, ele servesc la indicarea frecvenţei cu care diferite scoruri cu valoare mai mare sau mai mică sunt obţinute.

Deşi fiecare autor se opreşte asupra anumitor aspecte ale testului, putem rezuma în acest moment elementele de bază ale definiţiei testelor psihologice aşa cum reies din analiza celor enumerate.

Definirea elementelor	Explicaţii	Argumentare

Testele psihologice sunt proceduri sistematice, obiective	Sunt caracterizate prin planificare, uniformitate, standardizare, şi minuţiozitate.	Testele trebuie să fie, în mod demonstrabil, obiective şi corecte pentru a fi folositoare.
Testele psihologice sunt eşantioane de comportament.	Sunt mici părţi dintr-un întreg mult mai mare.	Eşantionarea comportamentului este eficientă deoarece timpul disponibil este, în mod obişnuit, limitat.
Comportamentele selectate pentru teste sunt relevante pentru funcţionarea cognitivă sau afectivă a persoanei, sau pentru ambele.	Eşantioanele de comportament sunt selectate datorită semnificaţiei lor empirice şi psihologice.	Testele, spre deosebire de jocurile mentale, există pentru a fi utile; sunt instrumente. Diagnosticul sau puterea predictivă depinde de relevanţa eşantionului de comportament selectat.
Rezultatele la teste sunt evaluate şi scorate(transformate în scoruri- cantităţi).	Un anume sistem, de categorii sau numere, este aplicat rezultatelor la teste, în concordanţă cu reguli prestabilite.	Nu trebuie să existe nici o îndoială, nici o discuţie, legată de ceea ce reprezintă rezultatele la teste.
Pentru a evalua rezultatele testului (a afla semnificaţia lor) este necesar să avem standarde bazate pe date empirice.	Trebuie să existe un mod de a aplica rezultatelor o măsură comună (un etalon) sau un criteriu comun.	Numai standardele utilizate pentru evaluarea rezultatelor testului oferă singura semnificaţie pe care acestea o pot avea.

Tabel. Elemente de bază ale definiţiei testelor psihologice

Alte noţiuni care au legătură cu testul psihologic

Scoruri- valori numerice care reprezintă evaluarea răspunsurilor la un test, după reguli precizate

Item- un element al testului psihologic constituit dintr-o situaţie stimul care solicită un răspuns ce poate fi evaluat separat de restul testului

Scală- în psihometrie sensul cel mai utilizat este -un set de itemi care se referă la acelaşi aspect- un atribut psihic pe care testul îl măsoară sau un comportament pe care testul îl prezice, spre exemplu scala de Depresivitate a Chestionarului de personalitate FPI- Freiburg;

- uneori noţiunea de scală se referă la un întreg test, constituit din mai multe părţi- spre exemplu Scala de inteligenţă Stanford-Binet;
- mai multe subteste care se referă la acelaşi aspect- spre exemplu- scalele verbale ale testului de inteligenţă Wechsler;
- un instrument separat destinat să măsoare un singur aspect, spre exemplu scala de Locus of control- Intern- Extern a lui Julian B. Rotter;
- sistemul numeric utilizat pentru a culege răspunsurile subiectilor la un test, spre exemplu o scală de la 1 la 5, unde 1 semnifică- Acord puternic şi 5 semnifică- Dezacord puternic

Baterie de teste- este un alt termen des utilizat în titlurile testelor sau în lucrările despre psihodiagnoză. O baterie de teste este un grup de câteva teste, sau subteste, care sunt administrate în acelaşi timp unei persoane sau unui grup de persoane. Când mai multe teste fac parte din acelaşi pachet comercializat de către autor pentru a fi utilizate într-un scop specific, de obicei cuvântul baterie apare în titlu şi întreg grupul de teste este considerat ca un singur instrument, în ansamblu.

Termenul de baterie de teste mai este utilizat pentru a desemna orice selecţie specifică de teste făcută de către un psiholog pentru a fi utilizate cu un client, pentru a da un răspuns la o problemă specifică, de obicei în scopuri diagnostice, sau de selecţie.

CLASIFICĂRI ALE TESTELOR

Având în vedere că există extrem de multe teste, problema clasificării lor şi, legat de aceasta, a criteriilor de clasificare devine importantă. Desigur, mai mulţi autori au încercat să realizeze clasificări, însă noi ne vom referi la patru autori români: Horia Pitariu, Ursula Şchiopu, Nicolae Mitrofan şi Monica Albu. Pentru uşurinţa vizualizării vom reda clasificările sub formă de tabel.

Clasificarea testelor după Horia Pitariu (1993)

Criterii	**Subcriterii**	**Categorii de teste**
Conţinutul testului	• După sarcină	• teste verbale- orale sau scrise • teste nonverbale- manipulare de aparate şi piese sau cu ajutorul calculatorului • de performanţă
	• După modul de procesare	• teste intelective- aptitudini, cunoştinţe,situaţionale,probe de lucru • teste nonintelective (chestionare/inventare de personalitate)
Administrarea testului	• Exigenţe legate de *eficacitatea* examenului psihologic	• teste individuale • teste de grup sau colective
	• Exigenţe legate de *timpul de examinare*.	• teste de viteză- cu limită de timp • teste de randament, care pot fi: cu limită de timp, cu timp nelimitat
	• Standardizare	• Teste standardizate • Teste nestandardizate
Cotarea testului	• Obiectivă	Teste obiective- performanţa se apreciază după o procedură fixă, impersonală
	• Subiectivă	Teste proiective- prezintă în cotare o doză mare de subiectivism

Clasificarea testelor după Ursula Şchiopu(2003)

Criterii	**Categorii**	**Exemple**
Sarcinile psihodiagnozei,, scopul	• cunoştinţe(pedagogice) • inteligenţă; • aptitudini; • senzorio-motorii; • atenţie • memorie • personalitate • sociabilitate • creativitate • temperament • interese	De inteligenţă: • Generală • Tehnică • Abstractă • Socială • Emoţională • kinestezică. testele de atenţie • concentrarea • stabilitatea • mobilitatea • distributivitatea Memoria cifrelor, a

		figurilor; Memoria topografică De aptitudini speciale • tehnice • numerice • matematice • artistice • verbale • pedagogice • organizatorice
Diferenţele semnificative de strategie incluse în structura de ansamblu a bateriei	Niciun fel de legătură între testele unei baterii	Măsoară în mod independent caracteristici psihice total diferite
	Între teste poate să existe o corelaţie pozitivă, cu valoarea coeficientului de corelaţie ce tinde către valoarea 1	când măsoară aceeaşi însuşire psihică.
	Mai multe teste pot fi saturate diferit în ceea ce priveşte fiecare dintre aceşti factori comuni	
Tipul de material administrat în teste şi caracteristicile acestuia	Testele „creion-hârtie”	Orice test la care răspunsul se dă pe o foaie de hârtie
	Testele- aparate	pompa Schultz, strungul Lahy etc.
	Testele ce folosesc diferite materiale	cuburile Kohs, asamblarea de imagini etc.
	Testele verbale	Ex. NEO PI R
	Testele nonverbale	Ex. NPQ
Diferenţe în ceea ce priveşte forma de administrare a testelor	Individuale	Un singur subiect, ex. Stanford –Binet, Wechsler, Szondi
	Colective	Un grup de subiecţi simultan(pot fi aplicate şi individual)
Caracteristicile tipurilor de răspunsuri solicitate	Cu răspuns la alegere	Ex. DA-NU, A-F, pe o scala de la 1 la 5 etc.
	Cu un singur răspuns bun	Mai ales de inteligenţă şi aptitudini
	Cu răspuns liber	Mai ales testele proiective, ex.Asociativ- verbal, Rorschach, TAT
Tipurile de răspunsuri primite	Sintetice, rezultă un coeficient global, sau un tip	Ex. IQ, EQ etc., tipul MBTI
	Analitice, rezultă un profil	Mai ales inventarele de personalitate

Un alt autor român, Nicolae Mitrofan(2007) observă, legat de aceste clasificări faptul că:

„În niciuna nu sunt luate în atenţie testele care sunt oferite nu neapărat psihologilor specialişti, ci nespecialiştilor.”, şi mai propune încă un criteriu şi anume gradul de standardizare, diferenţiind trei categorii de teste:

1. teste standardizate, cele care pot fi aplicate numai de către psihologi;
2. teste semistandardizate, cele care sunt destinate altor categorii de beneficiari, cum ar fi: cadrele didactice, părinţii, personal medical, asistenţi sociali etc.
3. teste nestandardizate, cele care se folosesc mai mult pentru ocuparea timpului liber .

Monica Albu(1998) deşi aplică aceleaşi criterii ca şi Horia Pitariu adaugă câteva categorii diferite. Spre exemplu în ceea ce priveşte testele de personalitate(nonintelective), apar trei categorii:

- Chestionare de personalitate
- Tehnici proiective
- Teste obiective de personalitate

După modul în care variază proba de la o administrare la alta:

- Cu conţinut fix- majoritatea
- Cu conţinut variabil
  - cele idiografice
  - cele adaptative – de aptitudini sau cunoştinţe, şi cele de personalitate

După modul de interpretare a scorurilor la teste sunt:

- Teste normative
- Teste centrate pe criteriu
- Teste idiografice
- Teste ipsative

O să aprofundăm discuţia pentru anumite tipuri de teste, mai frecvent utilizate:

- Chestionarele de personalitate
- Testele proiective
- Testele obiective de personalitate
- Clasificarea în funcţie de modul în care variază proba de la o administrare la alta
- Clasificarea în funcţie de modul de cotare a testului şi de interpretare a scorurilor

4.3. CALITĂŢILE GENERALE ALE TESTULUI ŞI „ECHIPAMENTUL TESTULUI PSIHOLOGIC”

Pentru a decide alegerea unui anumit test pentru o situaţie concretă este necesar să se evalueze adecvarea lui la obiectivele examenului psihologic, adică în ce măsură testul este relevant pentru diagnoza sau predicţia comportamentului care constituie scopul specific al examenului respectiv. Există o evaluare practică a testului şi o evaluare tehnică. Privind factorii ce determină valoarea practică a testului putem enumera:

- Utilitatea
- Acceptabilitatea
- Sensibilitatea
- Caracterul direct
- Nonreactivitatea
- Caracterul compatibil

Utilitatea se referă la avantajul practic pe care îl putem obţine din utilizarea unui anume instrument(Gottman, Leiblum, 1974, apud Minulescu, M., 2003), măsura în care instrumentul ajută la rezolvarea unei probleme practice- intervenţie, feedback privind eficienţa intervenţiei, selecţie, promovare etc. Alte aspecte ale utilităţii sunt: scopul testului, lungimea, uşurinţa în scorare, uşurinţa în intrepretare.
Acceptabilitatea se referă la necesitatea ca instrumentul să poată avea sens pentru persoană- adică să fie potrivit la condiţia intelectuală şi emoţională a persoanei, să reflecte exact problema persoanei şi să fie privit pozitiv de subiect (să nu fie prea intruziv sau ofensiv).
Sensibilitatea se referă la capacitatea testului de a sesiza schimbările în timp. Mai ales în consiliere sau terapie este nevoie de un instrument capabil să sesizeze progresul persoanei.
Caracterul direct se referă la capacitatea testului de a reflecta comportamenztul real, gândurile sau sentimentele reale ale persoanei. Măsurile directe sunt considerate semne ale problemei, în timp ce măsurile indirecte sunt considerate simboluri ale problemei, ceea ce necesită inferenţe, adică posibilităţi de eroare. Testele de performanţă sunt considerate directe, în timp ce tehnicile proiective sunt considerate indirecte. Toate instrumentele se situează pe o dimensiune ce uneşte aceste două extreme. Atunci când se pune accent pe fidelitatea datelor caracterul direct al testului este decisiv.
Nonreactivitatea se referă la calitatea testului de a nu provoca o anumită reacţie subiectului, influenţând astfel răspunsul. Această calitate are legătură şi cu caracterul nonintruziv al conţinutului, şi cu dezirabilitatea acestora. Sunt indicate instrumentele care nu determină reactivitatea subiectului.
Adecvarea generală a instrumentului este criteriul final, adică măsura în care acesta este compatibil pentru o evaluare fără complicaţii; instrumentele prea lungi sau prea complicate ca scorare pot fi utile ca informaţie dar nepractice datorită consumului de timp. Adecvarea se referă desigur şi la aspecte mai tehnice: fidelitatea şi validitatea informaţiei, obiectivitatea datelor pe care se poate întemeia o decizie.

Privind valoarea tehnică a testului mai putem cita pe Lawrence M. Rudner(1994) "Questions To Ask When Evaluating Tests" (Practical Assessment, Research & Evaluation), care formulează un ghid al evaluării unui test care cuprinde următoarele întrebări, pe domenii de interes pentru cel care face selecţia:

1. Care este utilizarea ce se intenţionează pentru test? Ce interpretări simte autorul testului că sunt adecvate? Sunt identificate utilizări inadecvate?
2. Pentru ce anume este destinat testul? Care este baza pentru a considera că testul se aplică subiectului/subiecţilor dumneavoastră?

Există eşantioane potrivite pentru validare şi normare?

1. Cum au fost alese eşantioanele utilizate în testarea pilot, în validare şi normare? Cum este acest eşantion relationat cu populaţia dumneavoastră? Sunt ratele de participare potrivite?
2. A fost eşantionul destul de mare pentru a dezvolta estimări stabile cu fluctuaţii minime datorate erorilor de eşantionare? Există afirmaţii despre subgrupuri, sunt de ajuns de mulţi subiecţi în fiecare subgrup?
3. Dificultăţile de măsurare ale testului şi ale criteriului(dacă sunt) furnizează o bază adecvată pentru validarea şi normarea instrumentului? Sunt suficiente variaţii în scorurile testelor?

Despre dovezile fidelităţii

Testul este suficient de fidel pentru a permite estimări stabile ale nivelurilor de abilitate ale indivizilor în grupul ţintă.

Interesul cel mai mare îl prezintă consistenţa internă datorată erorilor de eşantionare a conţinutului, de obicei sursa cea mai importantă de eroare.

1. Cum au fost calculate estimările fidelităţii? Au fost utilizate metode statistice potrivite?(exemplu- fidelitatea half-split nu trebuie calculată în testele de viteză)
2. Care sunt dovezile de fidelitate pentru grupuri de subiecţi diferiţi şi cum au fost calculate?

3. Este fidelitatea suficient de înaltă pentru a garanta utilizarea testului ca bază pentru decizia asupra indivizilor?
4. În ce grad grupurile utilizate pentru calculul fidelităţii sunt similare cu grupurile pe care va fi utilizat testul?

Despre dovezile validităţii relative la criteriu
Testul prezice adecvat performanţa în domeniul în care va fi folosit.

1. Ce măsură a criteriului a fost utilizată pentru a evalua validitatea? Care este raţiunea pentru care a fost aleasă această măsură?
2. Este adecvată distribuţia scorurilor la măsurătorile criteriului?
3. Care este acurateţea predictivă globală a testului? Cât de precise sunt predicţiile pentru indivizii ale căror scoruri sunt apropiate de scorul limită ?

Despre dovezile validităţii relative la conţinut

Gradul în care itemii testului sunt reprezentativi pentru abilităţile în domeniul de interes.

1. Există o imagine clară a universului abilităţilor reprezentate de test? Ce tip de cercetare a fost făcută pentru a determina conţinutul dorit al testului sau pentru evaluarea sa?
2. Care este compunerea echipelor de experţi utiliuzaţi în validarea conţinutului? Cum au fost judecăţile obţinute/ solicitate?
3. Cât de similar este acest conţinut cu cel pe care doriţi să-l testaţi?

Despre dovezile validităţii relative la construct

Testul măsoară constructul psihologic "corect"(dorit)

1. Este cadrul conceptual pentru fiecare construct testat clar şi bine fundamentat? Care este baza pe care s-a concluzionat că scopul testului este relaţionat cu acest construct?
2. Cadrul conceptual furnizează o baza pentru ipoteze testabile care privesc constructul? Sunt aceste ipoteze sprijinite de datele empirice?

Despre administrarea testului

Există instrucţiuni clare, potrivite şi detaliate care privesc procedurile de administrare a testului.

Toate specificaţiile de administrare a testului, inclusiv instructajul pentru subiecţi, liitele de timp, utilizarea unor materiale de referinţă şi calculatoare, iluminare, echipamente, aşezare, monitorizare, cerinţe legate de camera de testare, secvenţele de testare, ora potrivită etc. trebuie descrise în detaliu.

1. Persoana care administrează testul înţelege cu precizie ce anume se aşteaptă de la ea?
2. Procedurile de administrare a testului reproduc condiţiile în care testul a fost validat şi normat?

Despre raportul în urma testului

Metodele utilizate pentru a raporta rezultatele testului, incluzând scorurile scalate, rezultatele la subteste şi rezultatele combinate ale testului, sunt descrise în detaliu împreună cu raţiunea fiecărei metode(testul trebuie să fie util pentru a lua decizii).

1. Cum sunt raportate rezultatele testului?Sunt scalele utilizate pentru raportare de ajutor pentru o utilizare potrivită a testului?
2. Ce materiale şi resurse sunt disponibile pentru ajutor în interpretarea testului?

Bias- uri ale testului sau itemilor

Testul nu este afectat sau ofensiv privitor la rasă, sex, limba nativă, originea etnică, regiunea geografică sau alţi factori.

Este de aşteptat ca dezvoltatorii de teste să fie atenţi la caracteristicile demografice ale subiecţilor testării.

1. Au fost itemii analizaţi statistic pentru a descoperi eventuale bias-uri? Ce metode s-au folosit? Cum au fost itemii selectaţi pentru a fi incluşi în versiunea finală a testului?
2. A fost analizat validitatea diferenţială testului între grupuri? Cum a fost desfăşurată această analiză?
3. A fost determinată influenţa nivelului de cunoaştere a limbii asupra rezultatelor testului? Poate fi folosit testul cu vorbitori non-nativi ai limbii respective?

După Susana Urbina(2004) există anumite situaţii cînd nu este recomandabilă utilizarea testelor. **Zece motive pentru a nu utiliza un test psihologic**
Sunt multe motive şi multe situaţii în care utilizarea unui test psihologic nu este recomandabilă; lista care urmează prezintă pe cele mai relevante dintre ele. Cu câteva excepţii, testele psihologice nu ar trebui utilizate de fiecare dată când una dintre circumstanţele următoare apar:

1. Scopul testării este necunoscut sau neclar pentru utilizator.
2. Utilizatorul nu este complet familiarizat cu toată documentaţia testului şi pregătit practic în procedurile legate de test.
3. Utilizatorul nu cunoaşte unde vor ajunge rezultatele testului, sau cum vor fi utilizate, sau nu poate proteja folosirea lor.
4. Informaţia care poate fi obţinută prin testare este deja valabilă, sau poate fi obţinută mai eficient prin intermediul altor surse.
5. Subiectul testării nu este doritor sau capabil să coopereze pentru testare.
6. Subiectul testării poate experimenta consecinţe neplăcute datorită procesului de testare în sine.
7. Condiţiile de testare sunt inadecvate.
8. Materialele testului şi formatul acestuia sunt nepotrivite pentru vârsta, sexul, backgroundul cultural sau lingvistic, disabilităţilor, sau oricăror alte condiţii care pot invalida informaţiile oferite de test.
9. Normele testului sunt depăşite ca timp, inadecvate, sau inaplicabiel pentru subiect.
10. Documentaţia asupra fidelităţii şi validităţii scorurilor testului este inadecvată.

Manualul testului

Manualul testului trebuie să cuprindă, conform celor mai cunoscute standarde similare cu ale noastre:

1. Informaţia generală despre:

- Titlul, ediţia, formele
- Autorii
- Data publicării, manualele la dispoziţie, materialele suplimentare
- Timpul de administrare
- Costul pentru caiet şi 50 foi de răspuns

2. Descriere pe scurt a scopului/ naturii testului

    - Tipul general al testului- individual, de grup, de performanţă, baterie de aptitudini, inventar de interese etc.
    - Natura conţinutului: numeric, verbal, spaţial, psiho-motor
    - Subteste în cadrul testului, scale, scoruri separate pentru scale
    - Tipuri de itemi

3. Evaluarea practică

- Trăsături sau aspecte calitative ale materialelor testului, referitoare la:
    - Uşurinţa utilizării testului
    - Atractivitate atestului
    - Adecvarea faţă de subiecţii testaţi
- Aspecte privind modul de administrare- facilităţi de utilizare a computerului în administrare şi prelucrare etc.
- Proceduri de cotare- tipul de grilă, descrierea modului de cotare
- Calificări care se cer pentru administrarea testului: nivelul de formare cerut(C.P.I. cere un certificat pentru intrepretarea sa; testele de personalitate în general necesită formare);
- Validitatea de faţadă sau aspectul de faţadă- prezentarea testului pentru subiecţi;

4. Evaluări tehnice

- Normele- etaloanele testului- tipuri de norme- percentilele, scorurile standard
    - Eşantionul de normare- natura, mărimea, reprezentativitatea, proceduri pentru obţinerea eşantionului, criteriile pe baza cărora s-au stabilit subiecţii
    - Grupele de normare- vîrsta, sexul, educaţia, ocupaţia, regiunea din care provin subiecţii
- Fidelitatea- marja preciziei de măsurare- priveşte tipurile de fidelitate şi procedurile:
    - Tipuri de fidelitate- prin retestare, prin forme paralele, prin înjumătăţire testului, Kuder-Richardson, prin coeficientul Alpha- Cronbach
    - Alte informaţii procedurale- mărimea şi natura eşantionalor folosite pentru calculul fidelităţii, echivalenţele diferitelor forme de fidelitate, stabilitatea pe termen lung, eroarea standard a măsurării
- Validitatea- tipuri adecvate de proceduri de validare- de conţinut, de criteriu, predictivă sau concurentă, de construct
    - Proceduri specifice pentru măsurarea validităţii şi rezultatele lor, exprimate sub forma dovezilor validităţii
    - Mărimea şi natura eşantioanelor folosite

5. Revizii ale testului- direcţii, studii, date de cercetare
6. Evaluarea rezumativă a testului care priveşte:

- Aspectele pozitive majore sau slăbiciuni majore ale testului
- Testul ca instrument de diagnoză: dacă spre exemplu este inatacabil când se fac cercetări judiciare în legătură cu o diagnoză, deşi unii subiecţi atacă deciziile în urma testelor- se oferă credibilitate maximă testului

4.4. PROCEDURA DE SELECŢIE A UNUI TEST

Parafrazâd un personaj celebru al lui Shakespeare, Susana Urbina(2004) afirmă "To Use or Not to Use (Tests), That Is the (First) Question". Admiţând un răspuns afirmativ atunci mai rămâne o singură problemă, cum anume o facem. Mihaela Minulescu (2003) oferă un posibil ghid de decizie:

Cum selectăm un test?

1. Cadrul de referinţă care poate ghida alegerea unui instrument de măsură trebuie să răspundă la următorul set de întrebări(arbore decizional):

- Ce tip de evaluări sau decizii trebuie realizate?
- Ce tip de informaţie este necesară pentru a realiza evaluări şi decizii optime?
- De ce tip de informaţie dispunem deja?
- Cum şi de unde poate fi obţinut restul de informaţie?
- Ce metode şi/ sau instrumente avem la dispoziţie pentru a obţine această informaţie?
- Există şi/sau avem testele respective sau trebuie create?
- Cum trebuie evaluate testele sau instrumentele respective pentru a fi adecvate cerinţelor specifice?
- Ce alte criterii trebuie să utilizăm în alegerea lor?

2. Evaluarea testului sau instrumentului presupune compararea acestuia cu o serie de cerinţe practice şi tehnice, dintre care esenţiale sunt: fidelitatea- precizia măsurării, dovezi ale validităţii- capacitatea testului de a prezice performanţa sau de a evidenţia un diagnostic, interpretabilitatea şi predictibilitatea, toate acestea judecate în contextul concret al aplicării – scopul testării, populaţia pe care se aplică, condiţiile de testare.

Un posibil **checklist** pentru instrument poate cuprinde:

- Natura constructului
- Numărul de persoane ce trebuie testate
- Caracteristici ale examinatorului
- Caracteristici ale celui examinat
- Constrângeri practice
-

3. Criterii finale pentru decizia privind alegerea testului:

- Este testul destinat să măsoare comportamentul avut în vedere? Manualul său descrie scopurile testului; descrie adecvat comportamentele pe care le măsoară?

- Itemii testului pot măsura trăsăturile, obiectivele sau comportamentele ce trebuie evaluate?
- Este disponibilă informaţia privind diferitele aspecte referitoare la validitatea testului? Trebuie să determinăm dacă testul este valid pentru scopul specific, avut în vedere în prezent. În măsura în care testul cuprinde un număr limitat de itemi care să reprezinte direct obiectivele sau comportamentele de măsurat este important să stabilim că aceştia sunt într-un înalt grad reprezentativi pentru condiţia reală a celui testat. În aceeaşi măsură trebuie să verificăm, pornind de la faptul că testele folosesc constructe ipotetice care au fost definite operaţional de autorii lor, dacă există dovezi privind validitatea de construct a testului; în ce măsură rezultatele testului pot descrie competenţe într-o anumită aptitudine sau domeniu de conţinut.
- Este testul demn de încredere- fidel? Stabilitatea sau consistenţa testului în timp, a itemilor săi; cât de precis pot fi interpretate rezultatele? Dacă testele trebuie interpretate în raport cu persoane, este necesară şi utilă informaţia privind erorile standard de măsurare(un test cu SEM-eroarea standard a măsurării, nu este adecvat pentru diagnoza individuală şi nici în decizia privind plasamentul sau predicţii).
- Sunt furnizate informaţii suficiente pentru psiholog pentru a putea interpreta rezultatele? Etaloanele puse la dispoziţie sunt adecvate, persoana testată face parte din grupul de referinţă al etalonului disponibil? Există informaţie despre grupul de normare? Sunt specificate tehnici pentru a interpreta testul şi înţelesul diferitelor scale?
- Furnizează testul un feedback de interpretare pentru cel examinat? Putem furniza informaţii celui/celor testaţi legate de felul cum au performat la test? Poate fi înţeleasă informaţia?
- Este adecvat testul pentru cel examinat? Se iau în considerare nivelul de informaţie şi chiar de citire propriu-zisă pentru itemii testului. Dacă este prea dificil, persoana examinată poate avea dificultăţi de înţelegere a vocabularului, expresiilor sau exemplelor pe baza cărora răspunde. Intervin şi probleme mai concrete legate de mărimea literei, forma de prezentare şi maniera de a răspunde la test. Este adecvat testul vârstei sau condiţiei psihologice a celui testat?
- Este testul neinfluenţat de factori precum prejudecăţile sau deprinderile? Scopul este să avem itemi care sunt relevanţi şi pot fi înţeleşi în sensul lor propriu, nu figurat.
- Este testul, din perspectiva aspectelor practice, util? Problemele practice: costul, materiale suplimentare, timp de administrare, de interpretare etc. pot fi uneori decisive în selecţia unei probe pentru un program de testare.

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition),*Routledge, London and New York,
Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului,* Ed. Fundatiei Romania de Maine, Bucuresti,
Mitrofan Niculae- Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare,* Editura Polirom, Iaşi
Urbina Susana, *2004- Essential of psychological testing,* John Wiley & Sons, Inc. Hoboken, New Jersey,

5. MĂSURAREA ÎN PSIHODIAGNOZĂ

-SINTEZĂ-

5.1. MĂSURAREA ÎN PSIHOLOGIE ŞI TIPURI DE SCALE DE MĂSURARE

Măsurarea poate fi definită ca un proces sistematic de atribuire a unui număr unei anume variabile, respectiv „cuantificarea" aspectului psihologic. Variabila psihologică poate fi comportamentul clientului, sentimentele, atitudinile, scopurile sau rezultatul unui tratament, concepte teoretice ca estimarea de sine sau imaginea de sine. Numărul atribuit reprezintă caracteristica variabilei respective. Există mai multe definiţii ale măsurării(Albu, M., 1998).

Majoritatea acestora sunt de forma următoare:

"măsurarea constă în atribuirea de *x* lui *y* în conformitate cu z".

În funcţie de autor:

x reprezintă ***"numere", "valori", "scoruri" sau "simboluri abstracte";***

y este înlocuit cu ***"obiecte", "evenimente", "lucruri", "situaţii", "indivizi", "comportamente", "observaţii", "atribute", "proprietăţi" sau "răspunsuri";***

Z primeşte semnificaţia de "***regulă specifică" sau "regulă de reprezentare***"(Michell, 1997).

Prin măsurare a unei caracteristici la o mulţime de obiecte înţelegem atribuirea de numere obiectelor studiate, conform unei reguli bine precizate, prin care se urmăreşte ca anumite relaţii existente între obiecte, în privinţa caracteristicii respective, să existe şi între numerele atribuite lor. Mai precis, dacă în mulţimea obiectelor este definită o anumită relaţie (de exemplu, o relaţie de ordine), atunci trebuie să existe o relaţie similară în mulţimea măsurilor elementelor, iar dacă două obiecte sunt în relaţia considerată atunci măsurile lor trebuie să fie în relaţia similară ei. De pildă, dacă din punctul de vedere al caracteristicii cercetate un obiect *X* este superior unui obiect *Y*, atunci măsura obiectului X trebuie să fie mai mare decât măsura obiectului *Y*. Înseamnă, deci, că măsurarea stabileşte un izomorfism între mulţimea obiectelor şi mulţimea măsurilor obiectelor. în asemenea condiţii se pot efectua diverse operaţii asupra măsurilor obiectelor, iar concluziile desprinse se extind asupra obiectelor.

Din punct de vedere matematic, măsurarea este o **funcţie** care ataşează fiecărui element din mulţimea de obiecte (mulţimea de definiţie a funcţiei), conform unor anumite reguli, un număr şi numai unul (din mulţimea în care funcţia ia valori), numit "măsura elementului".

Fiecare sistem de reguli impus defineşte câte un tip de măsurare sau câte un tip de scală. În general sunt utilizate patru tipuri de scale de măsură. În ordinea crescătoare a preciziei lor, acestea sunt:

- **scala nominală**- precizie minimă, prelucrări statistice minime, concluzii puţine;
- **scala ordinală**- precizie mai mare, ordonare;
- **scala de interval**- precizie suficientă, majoritatea prelucrărilor statistice, concluzii;
- **scala de raport**- precizie maximă, prelucrări statistice complexe, concluzii numeroase.

Regulile impuse pentru definirea unui tip de scală le includ, totdeauna, pe cele impuse pentru definirea scalelor aflate pe un nivel inferior. De exemplu, dacă s-a putut defini o scală de interval, atunci s-ar fi putut defini şi o scală ordinală sau una nominală.

REŢINEŢI

1. Alegerea unui anumit nivel de măsurare pentru o mulţime de obiecte este dependentă de numărul şi tipul relaţiilor existente între obiecte şi care prezintă interes pentru cercetător. Evident, aceste relaţii se referă la caracteristica analizată.

2. Prelucrările statistice care pot fi efectuate asupra măsurilor obiectelor pentru a cunoaşte mai bine obiectele sunt dependente de tipul de scală utilizat. Cu cât scala de măsură este mai precisă, cu atât prelucrările statistice permise sunt mai complexe şi conduc la concluzii mai numeroase.

Măsurarea pe o scală nominală(categorială):
Poate fi utilizată de fiecare dată când caracteristica studiată permite împărţirea elementelor mulţimii de obiecte în clase disjuncte două câte două. În fiecare clasă sunt grupate toate obiectele care pot fi considerate echivalente în privinţa acelei caracteristici.
Exemplu: categoria socio-profesională, sex(1 poate semnifica- sexul masculin, 2- sexul feminin, sau M – masculin şi F- feminin), naţionalitate, etc.
Regula impusă la măsurarea nominală este ca toate elementele dintr-o aceeaşi clasă de echivalenţă să aibă aceeaşi măsură, şi reciproc, o valoare a măsurii să apară numai la elemente echivalente între ele. O posibilitate de definire a unei scale nominale constă în atribuirea de numere distincte(alese arbitrar) claselor de obiecte, toate obiecte având astfel aceeaşi măsură- cea ataşată clasei. Măsura unui element este de fapt o „etichetă" atribuită elementului, pentru denumirea, identificarea sau clasificarea acestuia. Poate fi exprimată numeric sau prin cuvinte, deci nu se pot face nici un fel calcule matematice cu măsurile nominale. Prin măsura nominală se poate obţine o singură informaţie asupra elementelor mulţimii de obiecte- două obiecte pot sau nu să facă parte din aceeaşi clasă. De aceea este bazală.

Măsurarea pe o scală ordinală:
Impune ca între elementele mulţimii de obiecte să existe o ordine, dar cu permisiunea ca mai multe obiecte să ocupe acelaşi loc. Spre exemplu persoanele dintr-un compartiment de muncă al unei firme pot fi ordonate în funcţie de randamentul lor, de la cel mai eficient la cel mai puţin eficient. În acest fel acceptăm convenţia că un element aflat mai sus în clasament este „superior" unui alt element care se află mai jos în acelaşi clasament (este mai mare decât, este mai bun decât, este preferatul lui), dar nu avem nici o informaţie despre mărimea distanţei între cele două elemente. Dacă două elemente ocupă acelaşi loc în şirul ordonat, vom spune că sunt echivalente.
Reguli: este valabilă regula de la măsurarea nominală şi în plus dacă un element este superior altui element, atunci şi măsura sa va fi superioară(dar nu ştim cu cât).
Exemplu: calificativele pe care jucătorii unei echipe de fotbal le primesc după un meci- de la foarte bun, bun, mediu, slab, foarte slab- indică o scală ordinală; severitatea depresiei descrisă în DSM IV R, şi în general severitatea simptomelor în clinică- indică o scală ordinală.
Scala este utilizată în mod curent dacă este cerută autoevaluarea- unei preferinţe, atitudini, comportament- ce pot fi descrise prin termeni de genul-„foarte puţin", „puţin", „mediu", „mult", „foarte mult". Nu se pot face operaţii matematice altele decât ordonarea- poziţia relativă, deoarece nu putem spune nimic despre distanţa dintre două elemente.

Măsurarea pe o scală de interval:
Poate fi utilizată atunci când există categorii diferite, un clasament şi în plus putem aprecia „distanţa" care desparte oricare două elemente în privinţa caracteristicii studiate.

Reguli valabile: sunt valabile regulile de la celelalte scale- categorii diferite, ordonarea, şi în plus putem aprecia distanţa dintre două elemente, care este egală între oricare două elemente consecutive- echidistanţa. Ceea ce lipseşte este faptul că nu avem un punct de plecare egal cu 0, adică absenţa totală a caracteristicii respective. Calculele matematice ce pot fi făcute sunt- adunare, înmulţire, împărţire, scădere. Lipsind valoarea lui zero absolut nu putem cu adevărat să cunoaştem semnificaţia multiplicării sau împărţirii scorurilor, deoarece un scor de 25 la un chestionar s-ar putea să nu reflecte jumătate din scorul 50. Putem spune că scorul 50 este mai mare cu 25 decât 25, dar nu putem afirma că este de două ori mai bun.
Exemplu: scala de temperatură- are intervale egale- echidistante, are şi valoarea 0, dar valoarea 0 nu înseamnă absenţa temperaturii. Diferenţa între 10 grade şi 15 grade este aceeaşi ca între 30 de grade şi 35 de grade, dar nu putem afirma nimic despre cât de mare este 15 grade în comparaţie cu 10 grade.
Alt exemplu- calendarele. Există un punct zero dar el este fixat arbitrar.
Normele testelor se bazează scala de interval. Putem compara mai multe caracteristici ale unui obiect între ele, măsurate pe o scală de interval, chiar dacă fiecare dintre măsurători furnizează valori în alt interval. Pentru aceasta se calculează scorurile standard, prin transformarea scorurilor brute.

Măsurarea pe o scală de raport:
Poate fi utilizată când avem categorii, ordonare, intervale egale şi în plus se poate aprecia obiectiv când un element este lipsit de caracteristica cercetată. Atunci când caracteristica măsurată reprezintă spaţiu (lungime, suprafaţă, volum), masă, sau timp trecut de la un eveniment fixat drept origine(durata de viaţă, durata de la un anumit eveniment precizat şi pâna la apariţia altui eveniment precizat, timpul de reacţie etc.) se poate stabili în mod obiectiv ce însemnă absenţa caracteristicii(adică punctul zero absolut). Pentru asemenea variabile se pot face măsurători de raport. În acest caz se poate deduce decâte ori o măsură este mai mare sau mai mică decât o altă măsură.
Exemplu: înălţimea de 1,80 m este de două ori mai mare decât 90 cm, deci persoana este de două ori mai înaltă; un autovehicul care se deplasează cu 100 km la oră merge de două ori mai repede decât una care are 50 km la oră; vârsta de 40 de ani este de două ori mai mare decât cea de 20 de ani.
Observaţie: dacă valorile metrice se transformă spre exemplu în trei grupe: sub 20 de ani, între 20 de ani şi 40 de ani, şi peste 40 de ani- atunci scala de raport se transformă într-o scală cu precizie mai mică- o scală ordinală. Dacă se transformă în cote standard valorile unei scale de raport- atunci noua scală va fi de interval. Măsurile de tip proporţie sunt relativ rare în domeniul ştiinţelor comportamentale.

Scalele diferă deci între ele prin patru caracteristici: exclusivitatea, ordonarea, echivalenţa, caracterul absolut.
Scalele de interval şi de proporţie se mai numesc şi **metrice(**variabilele obţinute se mai numesc **scoruri)**, iar cele nominale şi ordinale **nonmetrice.**

Alegerea unui tip de scală pentru măsurarea unei variabile trebuie să aibă în vedere, în primul rând, tipul de prelucrări statistice care ar trebui efectuate asupra variabilei. Dacă relaţiile cunoscute între elementele mulţimii de obiecte nu permit măsurarea variabilei pe scala care ar fi necesară calculelor statistice se va realiza o măsurare pe o scală inferioară ca precizie dar cât mai apropiată.

Măsurarea cantitativă în psihologie include trei tipuri de statistici, şi anume:
1. statistici descriptive;
2. statistici inferenţiale;
3. statistici multivariate.

1. Statisticile descriptive: grafice, modul de prezentare, de manipulare simplă a datelor (media, mediana, frecvenţa).
2. Statisticile inferenţiale: permit formularea unor concluzii referitoare la unul sau mai mulţi indivizi, bazându-se pe datele de eşantion, de lot. de grup de indivizi. Acest tip de statistici se referă la corelaţii, varianţa comportamentală, capacitatea de a judeca diferenţele dintre indivizi sau dintre grupurile de indivizi.
3. Statisticile multivariate: sunt utilizate atunci când punem în comparaţie două sau mai multe caracteristici măsurate pe un grup de indivizi, relaţiile dintre interese, valori şi gradul de diferenţiere dintre diferitele curbe variaţionale, corelaţii multiple, calcularea regresiilor, calcularea de tip factorial.

5.2.CALITĂŢILE METROLOGICE ALE TESTELOR PSIHOLOGICE

Testele psihologice constituie instrumente verificate ale căror principale calităţi metrologice sunt:

- fidelitatea
- validitatea
- standardizarea.

Fidelitatea unui test se referă la consistenţa rezultatelor obţinute de acelaşi subiect examinat la retestarea cu o probă identică sau o formă echivalentă. O stabilitate perfectă în timp a rezultatelor este dificil de obţinut datorită variabilităţii comportamentale în funcţie de factori fiziologici, motivaţionali, educaţionali etc. Al doilea aspect al fidelităţii îl reprezintă consistenţa internă, coerenţa prin eliminarea factorilor redundanţei sau neomogeni cu conţinutul testului. Toţi itemii testului trebuie să conveargă spre obiectivul probei respective.
Examinarea individuală a (itemilor testului are în vedere: stabilirea procentului de reuşită, eliminându-se din test cei cu un procent de reuşită peste 80% şi cei cu procentaj sub 20%, şi a corelaţiei item - test. Purificarea internă a testului este o condiţie centrală metrologică.

Validitatea este înţeleasă azi ca un complex de măsurători care se referă la mai multe aspecte, centrate în principal pe problema în ce măsură testul măsoară ceea ce îşi propune să măsoare. Testul este util, deci are valoare diagnostică sau prognostică, în funcţie de posibilitatea sa reală de a evidenţia gradul în care este prezentă o anumită abilitate sau trăsătură măsurată, precum şi dacă este capabil să anticipeze realizarea unei performanţe într-o anumită activitate. Aceste virtuţi specifice testului psihologic sunt condiţionate de validitatea sa. De exemplu, obiectivitatea testului este direct proporţională cu semnificaţia coeficientului de corelaţie (covariaţie) dintre test şi criteriul practic. Un test căruia nu îi cunoaştem validitatea constituie doar un instrument de cercetare, nu un instrument psihodiagnostic, echivalent unei ipoteze. Pentru determinarea validităţii sunt necesare două categorii de date: rezultatele la test şi rezultatele la un tip de măsurători pentru comportamentul avut în vedere de test, denumit criteriu. De exemplu, în selecţia profesională se întâlnesc două tipuri de validare: predictivă şi concurentă, fiecare utilizabilă în anumite condiţii şi având astfel o semnificaţie restrânsă de limitele acestor condiţii de aplicare.

Standardizarea se referă la posibilitatea de a obiectiva rezultatele unui subiect prin raportarea la performanţele obţinute la acelaşi test, în aceleaşi condiţii de administrare, de un eşantion reprezentativ de subiect. Interpretarea rezultatelor cere etalonarea testului, respectiv stabilirea unei scale care să permită reperarea facilă a performanţelor unui subiect în grupul total de performanţe al eşantionului reprezentativ, respectiv al grupului de persoane omogene subiectului sub raportul vârstei, nivelului educaţional, profesiunii etc. Rezultatele obţinute de acest grup de persoane reprezentative pentru condiţia subiectului testat se repartizează, de regulă, sub forma curbei lui Gauss, ceea ce permite calcularea performanţelor medii şi a dispersiei acestora. Media aritmetică a performanţelor la test este indicele care caracterizează tendinţa centrală a performanţelor acestor persoane, iar studiul variabilităţii rezultatelor (deviaţia standard sau abaterea standard) permite aprecierea specificului comportamentului acelui grup de persoane (comparativ cu un alt grup, cu o aceeaşi medie a performanţelor).

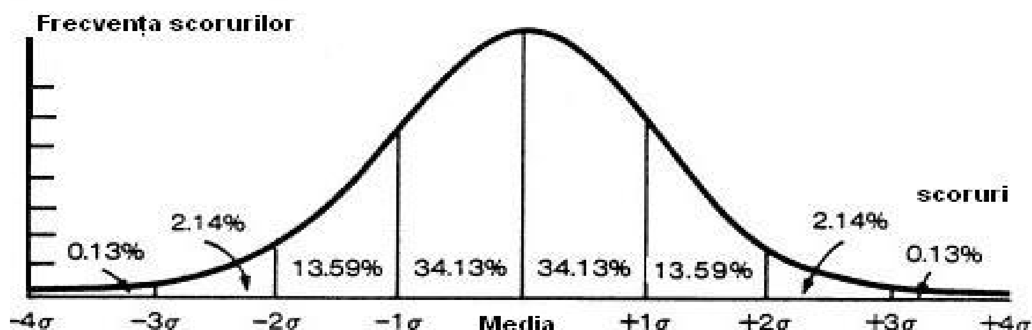
Interpretarea deviaţiei standard se realizează în funcţie de proprietăţile distribuţiei normale: o abatere în plus şi în minus faţă de medie acoperă relativ 68,26% din rezultate; două abateri standard, 95,44%; 3 abateri standard acoperă virtual totalitatea rezultatelor, 99,72%.

Câteva dintre proprietăţile curbei lui Gauss

Atunci când măsurăm dimensiuni ale personalităţii sau performanţe se observă, în general, o tendinţă ca frecvenţa rezultatelor obţinute să se distribuie într-un grafic ce ia forma unui clopot. Această tendinţă reprezentată grafic ca un clopot se numeşte curba probabilităţii normale sau curba întâmplării deoarece descrie distribuţia cea mai probabilă a frecvenţelor unor evenimente aleatorii(este valabilă în primul rând în domeniul ştiinţelor naturii). Această curbă are importanţă statistică datorită

faptului că aproximează repartiţia normală(în sensul că dacă măsurătoarea s-ar face pe un număr infinit de cazuri- atunci rezultatul ar fi exact) a unor măsurători efectuate pe un eşantion reprezentativ.

Specific curbei lui Gauss este simetria caracterizată prin aglomerarea spre centru a frecvenţelor cele mai mari şi scăderea lor treptată spre extremităţi. Spre exemplu- dacă testăm elevii privind cunoştinţele de istorie, printr-un test de cunoştinţe, vom observa că notele se vor distribui conform curbei lui Gauss.



Pot însă să apară şi excepţii, caracterizate prin abateri de la simetria distribuţiei. Atfel, dacă testul este prea uşor atunci vor fi foarte multe rezultate foarte bune, mai puţine la medii şi slabe şi în acest fel curba va fi înclinată spre dreapta- spre performanţele bune. Dacă apar două vârfuri sau două clopote, atunci testul a pus în evidenţa două grupuri de elevi, foarte diferite. Dacă testul este prea dificil atunci rezultatele se vor aglomera spre stânga curbei, în zona performanţelor slabe, mai puţine la medie şi bune (este înclinată spre stânga).

Curba lui Gauss stă la baza tuturor măsurătorilor biologice, psihologice, sociale etc., pe ea fiind clădită metoda statistică. Matematicianul german C.Fr. Gauss a avut ideea de a studia termenii dezvoltării binomiale pe cale geometrică, iar în domeniul măsurării diferenţelor umane metoda a fost aplicată de Sir Francisc Galton. În 1869 el prezenta în lucrarea sa Hereditary Genius, un sistem de clasificare a indivizilor după aptitudinile pe care le posedă(aptitudine- dimensiune a personalităţii care-i asigură unui individ reuşita într-o activitate). El sublinia că persoanele superior dotate sunt extrem de rare(1 la 4000), şi că toate aptitudinile umane sunt distribuite potrivit curbei lui Gauss, în consecinţă indivizii pot fi clasificaţi în funcţie de proprietăţile cunoscute ale distribuţiei normale. În timp s-a constatat că nu întotdeauna curba lui Gauss este utilă, pentru că nu întotdeauna măsurătorile au o distribuţie normală.

În practică, pentru a putea interpreta semnificaţia reală a rezultatelor unui subiect este necesar să le raportăm la rezultatele unei colectivităţi de referinţă, respectiv la un **etalon sau tabel normativ**.

Sistemele de etalonare cele mai utilizate în psihologia aplicată sunt: cvartilele, decilele. clasele normalizate (în 5,7,9 şi 11 clase). Procedura practică de elaborare a etalonului este în funcţie de distribuţia de frecvenţă a cotelor brute. De exemplu, **pentru distribuţii asimetrice** se preferă etaloanele în cvartile(4 clase), decile (10 clase) sau centile (100 clase). **Pentru distribuţii normale sunt adecvate etaloanele în clase normalizate.**

Colectivitatea de referinţă trebuie să constituie un **eşantion reprezentativ** din populaţia avută în vedere, iar etaloanele se raportează astfel la criterii precum: mediul din care provine populaţia (rural, urban), zona geografică, vârsta, sexul, nivelul educaţional. Se pot construi etaloane generale, cum este acela care are în vedere întreaga populaţie a unei ţări, culturi, spaţiu geografic larg etc, sau etaloane locale, a căror utilitate este restrânsă la aria particulară de referinţă.

Compararea rezultatelor la diferite teste este posibilă doar dacă acestea au la bază aceeaşi scară de notare. Nu se pot compara sau combina rezultatele exprimate în cote brute ale unui test de performanţă mnezică cu cele ale unui test de creativitate. Combinarea şi compararea cotelor nivelelor de performanţă, respectiv a gradului de semnificaţie a performanţei subiectului sunt posibile doar dacă le convertim în unităţi

standardizate, cote standard şi cote standardizate. In psihologie sunt preferate mai multe tipuri de cote standard: **variabilele normate Z(notele Z), variabilele T(notele T) şi H (Hull).**

Standardizarea testelor înseamnă, de asemenea, **standardizarea tehnicii de administrare, de cotare a răspunsurilor şi de interpretare a rezultatelor,** aspecte care au ca scop uniformizarea condiţiilor în care are loc examenul psihologic.

Toate aceste aspecte ale standardizării au în vedere **obiectivitatea,** respectiv eliminarea într-o măsură semnificativă a subiectivităţii examenului psihologic. Administrarea trebuie realizată în anumite condiţii obligatorii şi cu un anumit tip de instructaj, acelaşi pentru toţi subiecţii testaţi. Dintre cerinţe, semnificative sunt cele care privesc sala şi condiţiile de lucru, dar şi setul de cerinţe privitor la instructajul precis, clar, la timpul standard de exemplificare şi de testare propriu-zisă. Cu observaţia că, în examinarea psihologică, chiar şi în condiţiile examenului psihometric, se cere un anumit grad de supleţe a comportamentului celui care examinează, rigiditatea, respectiv standardizarea exagerată fiind la fel de negativă în consecinţe ca şi lipsa de standardizare. Acest lucru este semnificativ legat de faptul că, dintre factorii de influenţă, comportamentul examinatorului este doar un set de stimuli. Impactul acestui set de stimuli depinde însă şi de starea şi condiţia psihică a subiectului, respectiv, acelaşi comportament al celui care examinează(aceeaşi mimică, aceeaşi conduită verbală etc.) primeşte semnificaţii deosebite de la un subiect la altul (Cronbach, 1970).

5.3. CONCEPTE-CHEIE IMPLICATE ÎN METROLOGIE: CONSTRUCT, DOMENIU DE CONŢINUT, GRUPE CONTRASTANTE, CRITERIU, VALIDITATEA DE ASPECT,VARIABILĂ MODERATOARE ŞI VARIABILĂ MEDIATOARE

Construct

Definiţie: un construct este. în genere, **o idee construită(un model)** de experţi pentru **a reprezenta** (rezuma) un grup de fenomene sau de obiecte şi pentru a fi utilizată într-un cadru ştiinţific (teoretic, metodologic sau aplicativ).

Proprietăţi: După Murphy şi Davidshofer (1991), orice construct are două proprietăţi:

- este o abstractizare a unor regularităţi din natură;
- nu este observabil direct, dar poate fi conectat cu entităţi sau evenimente concrete, observabile.

Constructele psihologice sunt evidenţiate, direct sau indirect, de **comportament** sau în urma conducerii unor experimente dedicate măsurării lor.

Un construct psihologic, arată J.F. Binning şi G.V. Barett (Silva, 1993), este **„o etichetă aplicată unei grupări de comportamente care covariază".**

Pentru *Standards for Educaţional and Psychological Tests* din 1985, constructul este **o „noţiune teoretică derivată din cercetare şi din alte experienţe, care a fost construită pentru a explica modele de comportament observabile"**. Constructele create pentru interpretarea rezultatelor testelor psihologice sunt identificate, de obicei, cu câte o clasă de răspunsuri (L. J. Cronbach, după Silva, 1993).

Structura constructului: unii autori consideră constructele ca fiind multidimensionale, formate din „componente" independente unele de altele, numite uneori „faţete" şi recomandă utilizarea analizei factoriale ca metodă pentru studierea acestora.

De exemplu, în modelul elaborat de Guilford, constructul „creativitate" are drept componente: fluiditatea,flexibilitatea, originalitatea, elaborarea, senzitivitatea faţă de probleme şi redefinirea (Landau, 1 979).

Prezenţa componentelor ridică probleme în studiul relaţiei dintre construct, privit ca variabilă independentă, şi o variabilă dependentă prezisă de construct. O posibilitate de rezolvare o reprezintă cotarea separată a componentelor, cumularea scorurilor şi cercetarea relaţiei dintre scorul total obţinut şi valorile variabilei dependente. Dar, uneori, componentele ar putea să prezică mai bine variabila

dependentă decât constructul general. In asemenea situaţii se studiază separat relaţia dintre fiecare componentă şi variabila dependentă (Ozer, Reise, 1994).

Construct şi teorie psihologică: fiecare construct psihologic este fundamentat pe o teorie, care permite descrierea şi predicţia, a comportamentelor legate de el, în situaţii specifice.

Pentru elaborarea unui test care să îl măsoare, constructului i se asociază o variabilă cantitativă despre care se presupune că se află în relaţie crescătoare cu scorurile testului. Se formulează apoi o serie de afirmaţii referitoare la modul în care se comportă persoanele care au valori mari ale acestei variabile, comparativ cu cele care au valori scăzute.

De exemplu, Guilford, în teoria tridimensională a intelectului se referă la constructul gândire divergentă reprezentând creativitatea şi diferitele comportamente asociate abilităţilor din care este formată aceasta, diferenţiate pe unităţi, clase, relaţii.

Dorind să dezvolte un test de evaluare a motivaţiei indivizilor pentru propria realizare, McClelland şi colaboratorii săi au formulat următoarele descrieri de comportamente (Thomdike, Haagen, 1961):

• Persoanele care au o motivaţie puternică pentru realizare au rezultate şcolare bune.
• Motivaţia pentru realizare este mai mare dacă subiecţilor li se administrează în prealabil teste despre care li se spune că măsoară inteligenţa.
• Cei care au o motivaţie puternică pentru realizare completează mai mulţi itemi într-un test de viteză a cărui funcţie le-a fost explicată.
• Motivaţia pentru realizare este mai mare la copiii a căror familie le-a încurajat de timpuriu independenţa.

Fiecare dintre aceste descrieri este bazată pe un fel de „teorie a comportamentului uman". De exemplu, realizarea şcolară este văzută ca o combinaţie între aptitudinea şcolară şi efort. Dintre persoanele cu acelaşi nivel aptitudinal, cele care au o motivaţie mai marc pentru realizare depun un efort mai mare şi obţin rezultate şcolare mai bune (Thorndike, Haagen, 1961).

Construct şi comportamente: Comportamentele prin care autorul testului consideră că se manifestă constructul pot fi influenţate de diverse variabile, precum vârsta, sexul, educaţia primită etc. Pentru unele dintre acestea, legătura cu constructul poate corespunde unei situaţii reale, observată în populaţia căreia i se adresează testul şi acceptată de teoria care stă la baza constructului. Pentru altele, însă, ea denotă că aceste comportamente nu sunt caracteristice constructului. De asemenea, comportamentele prin care este descris constructul măsurat de test se pot asocia cu comportamente prin care se manifestă alte constructe psihologice, evaluate prin teste existente.

Descrierea constructului: în consecinţă, elaborarea unui test care să măsoare un anumit construct impune descrierea constructului în termeni comportamentali concreţi. Operaţia se numeşte explicarea constructului şi constă din următorii trei paşi (Murphy, Davidshofer, 1991):

(1)Se identifică acele comportamente care au legătură cuconstructul măsurat de test.
(2) Se identifică alte constructe şi pentru fiecare se decide dacă are sau nu legătură cu constructul măsurat de test.
(3) Pentru fiecare construct se alcătuieşte o listă de comportamente prin care acesta se manifestă. Pentru fiecare comportament, pe baza relaţiilor dintre constructe, se decide dacă are sau nu legătură cu constructul măsurat de test.

Rezultă astfel un sistem de relaţii care înglobează un set de constructe şi comportamente, denumit, de L. J. Cronbach, reţea nomologică („nomologică" în sensul că ea are la bază declaraţii exprimate sub formă de legi). Unii autori au folosit termenul de „reţea stohastică", susţinând că relaţiile dintre componentele reţelei sunt mai degrabă probabilistice decât de tip lege.

O descriere amănunţită a constructului trebuie să conţină şi relaţiile acestuia cu cât mai multe variabile fiziologice, psihologice etc.

De exemplu, dacă se urmăreşte realizarea unui test pentru evaluarea inteligenţei, se pot avea în vedere următoarele relaţii ale constructului „inteligenţă" cu alte variabile (Gronlund, 1967; Kline, 1993):

Scorurile testului cresc cu vârsta (se presupune că inteligenţa creşte cu vârsta aproximativ până la 16 ani).

- Scorurile testului prezic succesul şcolar.

- Scorurile testului concordă cu aprecierile făcute de profesori despre inteligenţa elevilor.
- Scorurile testului se corelează pozitiv cu scorurile unor teste aşa-numite „de inteligenţă".
- Dacă se efectuează analiza factorială a acestui test şi a altor teste care măsoară aptitudini, se va constata că testul de inteligenţă are o saturaţie mare în primul factor general.
- Scorurile testului discriminează între un grup de copii handicapaţi mintal şi unul format din copii capabili de performanţe superioare.
- Scorurile testului sunt influenţate doar puţin de procesul instrucţional urmat de subiect.
- Scorurile testului discriminează între diverse grupe ocupaţionale.

Unele profesii, precum cea de matematician sau de funcţionar superior, necesită o inteligenţă superioară, în timp ce altele pol fi exercitate şi de persoane cu o inteligenţă redusă, fiecare dintre afirmaţiile formulate în legătură cu constructul pe care testul trebuie să îl măsoare urmează să fie verificată experimental. Dacă unele dintre ele nu se confirmă, trebuie revăzută teoria care stă la baza constructului. pentru a sesiza eventualele omisiuni sau neconcordanţe.

Domeniul de conţinut al unui test

Definiţie: prin domeniul de conţinut al unui test se înţelege **„mulţimea tuturor comportamentelor care pot fi utilizate pentru a măsura atributul specific sau caracteristica la care se referă testul"** (Murphy, Davidshofer, 1991).

Definirea domeniului de conţinut este asemănătoare celei de definire a unui construct. Autorul testului, pe baza ideii pe care şi-a format-o despre ceea ce vrea să măsoare, reţine acele comportamente care presupune că ar fi manifestări ale constructului, respectiv, în cazul domeniului de conţinut, expresii ale performanţei sau ale cunoştinţelor subiecţilor.

Deosebirea dintre un construct şi un domeniu de conţinut constă numai în gradul de abstractizare implicat şi adesea este dificil de făcut distincţie între ele (Guion, 1986).

Descrierea domeniului de conţinut trebuie, în primul rând, să precizeze limitele acestuia. Apoi, pentru fiecare componentă se apreciază importanţa. In cazul testelor destinate să măsoare rezultatele unui program de instruire (cunoştinţe sau deprinderi), descrierea domeniului de conţinut al testului se bazează pe analiza conţinutului şi a obiectivelor instrucţionale ale cursului urmat de persoanele care vor fi testate. Se încearcă să se includă în domeniul de conţinut aspectele pe care mai mulţi experţi le consideră importante.

Exemplu: Un test de cunoştinţe de psihometrie, administrat studenţilor din anul II de la facultatea de psihologie, după predarea cursului de Psihodiagnostic, poate avea ca domeniu de conţinut „cunoaşterea noţiunilor de bază ale teoriei testelor, conform materiei predate la curs".

Descrierea domeniului de conţinut cuprinde:

a) limitele domeniului: se au în vedere cunoştinţele cuprinse în bibliografia indicată la curs (se anexează lista bibliografică);

b) componentele domeniului:

• probleme: analiza de itemi, validitatea testelor, fidelitatea testelor;

• aspecte tratate: definiţiile noţiunilor, metodele de cercetare, condiţiile de utilizare a metodelor de cercetare şi a formulelor de calcul;

c) importanţa relativă a componentelor domeniului: este indicată în tabel. Procentele se referă la timpul alocat fiecărei probleme în cadrul cursurilor predate studenţilor.

Problema	Aspectul tratat		
	Definiţii	Metode	Condiţii de utilizare
Analiza de itemi	12%	14%	8%
Validitatea testelor	11%	17%	5%
Fidelitatea testelor	6%	19%	8%

Grupele contrastante

Definiţie: Grupele contrastante sau grupele extreme reprezintă două loturi de subiecţi care, printr-o variabilă (sau mai multe), au valori diferite extreme, respectiv un grup are valori foarte mari, celălalt valori foarte mici.

Modul de formare a grupelor contrastante este dependent de numărul variabilelor alese şi de scala pe care acestea sunt măsurate. In cazul unei singure variabile dihotomice, cum este sexul, grupele contrastante corespund celor două valori ale variabilei.

Dacă pentru formarea grupelor contrastante se utilizează o singură variabilă, măsurată pe scală ordinală, de interval sau de raport, atunci se începe prin a ierarhiza persoanele pe baza valorilor acestei variabile. In prima grupă, denumită de obicei „grupa slabă'", vor fi incluse persoanele cu valorile cele mai mici ale variabilei, iar în cealaltă grupă, denumită „grupa bună", persoanele cu valorile cele mai mari. în literatura de specialitate există recomandarea ca cele două grupe să conţină câte 27% din totalul subiecţilor. Dar, din motive practice (se poate întâmpla ca mai multe persoane să aibă aceeaşi valoare a variabilei), se acceptă ca procentul celor incluşi în fiecare grupă să fie cuprins între 25% şi 33%.

Uneori, pentru formarea grupelor contrastante, este necesar să se ia în considerare mai multe variabile, fiecare fiind măsurată pe scală ordinală, de interval sau de raport. Aşa se întâmplă, de pildă, când grupele trebuie să conţină persoane „cu succes profesional" şi, respectiv, persoane „cu insucces profesional", iar performanţa profesională se exprimă prin valorile mai multor variabile. In asemenea situaţii, se procedează la ierarhizarea persoanelor după valorile fiecărei variabile şi la agregarea clasamentelor astfel formate. Din ierarhia rezultată se aleg cei aliaţi la extreme, ca şi în cazul descris mai sus.

Criteriul

Definiţie: Prin criteriu se înţelege „o variabilă pe care încercăm să o prevedem în general cu ajutorul testelor", sau „o măsură acceptată a comportamentului evaluat de test". Criteriul, în aprecierea personalului, este „o măsură a nivelului de performanţă exprimată în termeni cantitativi, bazată pe o descriere completă a muncii prestate".

Dimensiunile criteriului: variabila-criteriu poate fi unidimensională sau selecţia profesională a managerilor o variabilă criteriu poate fi constituită multidimensională, cu valori cantitative sau/şi calitative. De exemplu, în criteriul „eficienţă managerială" format din abilitatea de comunicare şi din capacitatea de organizare (apreciate prin câte un calificativ), din sociabilitate, persuasiune şi ambiţie (măsurate cu ajutorul unor scale, deci având valori numerice).

De exemplu, performanţa în muncă este, de asemenea, un criteriu multidimensional.

In compoziţia sa intră aspecte ce ţin de (Gilmer, 1971; Pitariu), 1994):

- cantitatea realizată în unitatea de timp (piese produse, cuvinte dactilografiate, itemi codificaţi etc);
- calitatea muncii (număr de piese rebutate, număr de erori de codificare etc); timpul de muncă neutilizat (absenţe, întârzieri, pauze neautorizate etc);
- stabilitatea în muncă (vechimea în muncă, schimbări de profesie, transferări, demisii etc);
- costul necesar pentru instruire (timpul de instruire pentru realizarea unei producţii standard, costul materialului stricat în timpul instruirii, viteza de avansare etc);
- cunoştinţele profesionale;
- satisfacţia în muncă (număr de revendicări, de sugestii, examinări psihiatrice etc);

- evaluările facute de superiori sau de colegi.

Se recurge adesea la verificarea corectitudinii unui test, în cadrul operaţiei de validare a testului, prin compararea scorurilor sale cu măsurile unui criteriu, atunci când testul trebuie să înlocuiască variabila-criteriu, fie din cauză că valorile variabilei-criteriu nu sunt cunoscute în momentul testării, lle pentru că administrarea testului este mai puţin costisitoare decât măsurarea variabilei-criteriu.

Condiţii: Pentru a face posibilă aprecierea corectă a calităţilor testului cercetat, criteriul trebuie să îndeplinească următoarele condiţii (adaptare după Greuter, Algera, 1989 şi Corsini, 1994):

1. Să fie relevant pentru activitatea sau caracteristica la care se referă, adică ordinea subiecţilor determinată pe baza valorilor criteriului să corespundă ordinii reale a acestora în privinţa succesului obţinut în activitatea respectivă sau a mărimii caracteristicii măsurate. Pentru aceasta este necesar ca variabila-criteriu să ia valori într-o mulţime total ordonată ori să realizeze o măsurare pe scală ordinală, de interval sau de raport. De obicei se preferă criteriile cu valori cantitative.
2. Să fie fidel, adică să concorde cu diverse evaluări ale performanţelor la care el se referă, făcute la momente diferite şi/sau cu alte instrumente de măsură, similare lui. Verificarea acestei condiţii se poate face cu ajutorul coeficientului de corelaţie.

Să fie practic, adică să nu coste mult şi să fie acceptabil pentru cei care doresc să îl utilizeze la luarea deciziilor. Pentru îndeplinirea acestei condiţii, este important ca în criteriu să fie înglobate principalele aspecte care sunt avute în vedere la aprecierea performanţei sau a mărimii caracteristicii măsurate. Identificarea acestora se poate face apelând la supervizori, respectiv la experţi.

4. Să fie exprimat în aceleaşi unităţi sau în unităţi comparabile pentru toate persoanele. De exemplu, dacă valorile variabilei-criteriu sunt dependente de vârsta subiecţilor, atunci aceste valori vor fî transformate în cote standardizate, pentru fiecare persoană fiind utilizate în formula de transformare media şi abaterea standard corespunzătoare grupei de vârstă din care aceasta face parte.
5. Atunci când criteriul este multidimensional, valoarea criteriului pentru fiecare persoană să reflecte atât importanţa diverselor componente, cât şi nivelul individului la fiecare dintre ele.

În *Criterii pentru analiza validităţii testelor care măsoară conslructe:* Anastasi (1954) enumera câteva criterii utilizate în practica psihologică pentru analizarea validităţii testelor care măsoară constructe. Acestea sunt:

a. Vârsta: se foloseşte drept criteriu pentru testele de inteligenţă, întrucât se consideră că scorurile acestora trebuie să crească pe parcursul copilăriei, până la maturitate. Nu este un criteriu bun pentru testele care măsoară funcţiuni psihice invariabile sau puţin variabile în timp, cum sunt variabilele de personalitate.

b. **Cunoştinţele şcolare,** exprimate prin note şcolare, rezultate lateste de cunoştinţe ori evaluări tăcute de profesori asupra „inteligenţei"elevilor: sunt adesea criterii pentru testele care măsoară aptitudinea şcolară.

c. **Performanţele la diverse programe de instruire specială** cum sunt cursurile de dactilografîe, programare, muzică etc: furnizează criterii pentru teste de aptitudini speciale.

d. **Performanţa profesională** serveşte uneori drept criteriu pentru teste de personalitate.

e. **Evaluările făcute de profesori, instructori sau superiori** sunt folosite mai mult pentru a obţine informaţii despre alte criterii, cum ar fi cunoştinţele şcolare, performanţa la diverse cursuri, reuşita profesională. Au rol de criteriu pentru unele teste care măsoară trăsături psihice precum anxietatea, onestitatea, originalitatea etc.

f. **Rezultatele altor teste** se utilizează ca şi criteriu atunci când se urmăreşte să se construiască un test care să măsoare acelaşi lucru ca şi criteriul, dar să fie mai uşor de administrat sau/şi de cotat, ori să fie mai ieftin. De exemplu, scala Stanford- B i n e t a servit drept criteriu pentru multe teste de inteligenţă.

g. **Grupele contrastante** reprezintă de obicei un criteriu multidimensional, care cuprinde aspecte diverse din viaţa zilnică. De exemplu, pentru un test de inteligenţă, grupele contrastante pot fi constituite dintr-o clasă de copii proveniţi de la o şcoală de debili mintali şi o clasă de elevi de aceeaşi vârstă dintr-o şcoală obişnuită. Criteriul pe baza căruia s-au format grupele este alcătuit din ansamblul de factori care au determinat cuprinderea copiilor în şcoala ajutătoare. Grupele contrastante se folosesc frecvent ca şi criteriu pentru testele de personalitate. De exemplu, pentru evaluarea unor trăsături legate de activitatea socială se poate forma o grupă de studenţi cu multiple preocupări extraşcolare şi una cu studenţi fără asemenea preocupări.

Clasificarea criteriilor: Criteriile utilizate în cercetările asupra comportamentului pot fi clasificate din două puncte de vedere, ambele permiţând plasarea criteriilor de-a lungul unei axe (Barclay, 1968). Astfel:

- după modul de obţinere a valorilor lor, la un pol se plasează criteriile *subiective* (valorile sunt furnizate de subiect), iar la celălalt, cele *obiective* (valorile sunt obţinute de la alte persoane sau cu ajutorul unor instrumente de măsură);
- după gradul în care valorile lor sunt dependente de contextul în care se află persoana evaluată, la un pol se

află criteriile *individuale* (valorile sunt independente de contextul în care se află persoana), iar la celălalt, cele *de grup* (comportamentele sunt puternic influenţate de grupul din care face parte persoana evaluată cu ajutorul criteriului).

De exemplu, criterii **subiective** individuale sunt afirmaţiile făcute de subiect despre nivelul său de aspiraţie şi despre valorile sale spirituale, în timp ce criterii subiective de grup sunt relatările subiectului despre relaţiile sale cu alte persoane, despre recompensele şi pedepsele pe care le-a primit.

Criterii **obiective** individuale sunt tehnicile Q-sort, iar dintre criteriile obiective de grup fac parte evaluările realizate de profesori şi observaţiile sistematice asupra comportamentului.

Mulţimea de criterii: Adesea, pentru un acelaşi test se găsesc mai multe criterii, fiecare referindu-se la un alt aspect. Aşa se întâmplă, de pildă, când testul trebuie să prevadă reuşita profesională.

In asemenea cazuri există două posibilităţi de a acţiona, pentru a le lua în considerare pe toate:

a) se combină criteriile într-un **criteriu compus** sau global; pe baza scorurilor criteriilor se calculează pentru fiecare persoană o cotă unică;

b) se reunesc criteriile într-un **criteriu multiplu:** fiecare criteriu are o valoare particulară şi este utilizat separat în aprecierea testului.

Prima variantă de lucru este mai convenabilă în privinţa prelucrării şi a interpretării datelor, întrucât fiecare persoană are un singur scor, care permite ierarhizarea candidaţilor. Pentru aplicarea sa, este necesar, însă, ca scorurile criteriilor să fie exprimate în aceleaşi unităţi de măsură: în scoruri standardizate, în ranguri, în unităţi de timp sau în costuri.

Exemplu: Pentru munca de editare a textelor la calculator se pot fixa două variabile criteriu:

- numărul de semne introduse într-o oră;
- numărul de erori comise la introducerea acestor semne.

Cu ajutorul acestor criterii se poate forma un criteriu global, care exprimă timpul necesar pentru a introduce corect un număr fixat de semne, N. Valoarea criteriului global, pentru fiecare persoană, se calculează însumând timpul în care sunt introduse cele N semne cu timpul necesar pentru depistarea erorilor comise şi corectarea lor. **Prin combinarea criteriilor** într-un criteriu global şi ierarhizarea persoanelor pe baza scorului astfel calculat se rezolvă o problemă de agregare a mai multor clasamente: ierarhiile de pornire (corespunzătoare scorurilor criteriilor) se înlocuiesc cu un singur clasament.

Nota globală: în practica selecţiei profesionale se obişnuieşte să se combine liniar, ponderat sau neponderat, scorurile criteriilor pentru a obţine o cotă globală. în privinţa modului de stabilire a ponderilor, părerile sunt diverse. De exemplu, H. Toops (Brogden, Taylor, 1949) consideră că ponderile trebuie să reflecte importanţa fiecărui criteriu, apreciată de specialişti, în timp ce M.L. Blum şi J.C. Naylor (1968) propun ca ponderile să fie proporţionale cu fidelitatea criteriilor.

Procedeele de agregare a clasamentelor care însumează scorurile ponderate ale criteriilor prezintă un neajuns atunci când ierarhiile corespunzătoare criteriilor nu concordă între ele. Astfel, este posibil ca pe primele locuri ale ierarhiei finale să ajungă persoane care au obţinut scoruri modeste la mai multe criterii, dacă ele s-au aflat pe locuri fruntaşe în clasamentele corespunzătoare criteriilor cărora li se atribuie ponderi mari.Acest inconvenient poate avea urmări grave atunci când criteriile nu au un caracter compensatoriu. Dacă existenţa unor cunoştinţe sau deprinderi este absolut necesară desfăşurării cu succes a activităţii pentru care se face selecţia şi nu poate fi compensată prin alte calităţi ale persoanei, de asemenea necesare postului, aceste cunoştinţe sau deprinderi trebuie să constituie un criteriu care nu va putea fi, sub nici o formă, inclus într-un criteriu global (Landy, 1985). în asemenea situaţii trebuie să se utilizeze un criteriu multiplu.

Validitatea de aspect a testelor

Definiţie: Orice test psihologic sau educaţional destinat utilizării în practică trebuie să fie perceput de cei examinaţi cu el ca fiind adecvat domeniului în care este aplicat. Această calitate a testelor este denumită validitate de aspect (C. 1. Moşier, după Nevo. 1993).

O definiţie operaţională a validităţii de aspect a fost formulată de 1 . Guttman şi S. Shye (Nevo, 1993) astfel: „Se face o măsurare a validităţii de aspect atunci când o persoană testată/un utilizator

neprofesionist/un individ interesat, utilizând o tehnică absolută /relativă apreciază un unu/ un test/o baterie de teste drept foarte potrivit/nepotrivit pentru utilizarea care îi interesează"

I rebuie observa) că aprecierea validităţii de aspect nu este realizată de experţi şi pentru măsurarea sa nu se folosesc metode psihometrice. Dacă aprecierea validităţii de aspect este făcută de persoanele testate, ele trebuie să îşi spună părerea imedial (după terminarea testului, când impresiile sunt încă proaspete. Validilalea de aspect nu este inclusă printre celelalte tipuri de validitate *(Stqndards for Educaţional and Psychological Tests,* 1974) şi nici nu le poate înlocui.

Variabile moderatoare, variabile mediatoare

Variabilă moderatoare

Definiţie: Noţiunea de variabilă moderatoare are mai multe accepţiuni.

In termeni generali, o variabilă moderatoare este o variabilă calitativă (de exemplu, sexul, rasa, clasa socială) sau cantitativă (de exemplu, vârsta, anxietatea), care afectează direcţia şi/sau tăria relaţiei dintre o variabilă independentă şi o variabilă dependentă (Baron, Kenny, 1986).

Unii autori înţeleg prin variabilă moderatoare „o interacţiune între mai multe variabile care influenţează corelaţiile uneia dintre variabile cu alte date" (Mischel, 1968).

Procedeu: Procedeul statistic utilizat pentru a verifica dacă o variabilă independentă acţionează sau nu ca o variabilă moderatoare este determinat de felul în care au fost măsurate variabila independentă şi variabila moderatoare. Se face distincţie între cazul când variabilele sunt categoriale, adică discrete, măsurate pe scală nominală sau ordinală, şi cel în care ele sunt continue, măsurate pe scală de interval sau de raport.

Atunci când variabila dependentă este continuă, sunt posibile următoarele situaţii (Hofstee, Smid, 1986; Baron, Kenny, 1986):

1. Variabila independentă şi variabila moderatoare sunt ambele categoriale.

în acest caz se aplică analiza de variantă bidimensională pentru a cerceta efectul variabilei independente şi al variabilei despre care se presupune că este moderatoare asupra variabilei dependente. Existenţa efectului moderator este indicată de prezenţa interdependenţei dintre variabila independentă şi cea moderatoare.

2. Una dintre cele două variabile (independentă şi moderatoare) este continuă, iar cealaltă categorială. în acest caz se determină ecuaţia de regresie a variabilei dependente faţă de variabila continuă, pentru fiecare valoare a variabilei categoriale. Dacă există un efect moderator, atunci dreptele de regresie nu sunt paralele.

3. Variabila moderatoare şi variabila independentă sunt ambele continue.

Variabilă mediatoare:

Definiţie: O variabilă *U* acţionează ca mediator între variabila independentă X şi cea dependentă *Y* dacă sunt îndeplinite următoarele condiţii (Baron, Kenny, 1986):

• variaţia valorilor variabilei independente provoacă variaţia valorilor variabilei mediatoare;

• variaţia valorilor variabilei mediatoare determină variaţia valorilor variabilei dependente;

• relaţia dintre variabila independentă şi cea dependentă este semnificativă, dar devine slabă sau chiar inexistentă atunci când sunt controlate relaţiile dintre variabila independentă şi cea mediatoare şi dintre variabila mediatoare şi cea dependentă.

Procedeul statistic: Procedeul aplicat pentru verificarea existenţei efectului mediator constă în construirea ecuaţiilor de regresie.

Medierea perfectă apare atunci când variabila mediatoare fiind controlată, variabila independentă nu are nici un efect asupra variabilei dependente.

în general, în cercetările psihologice, mediatorii sunt diverse caracteristici ale persoanelor care transformă variabila de intrare în variabila de ieşire.

Utilizare: Variabilele moderatoare sunt utilizate atunci când relaţia dintre variabila independentă şi cea dependentă este slabă, ea modificându-se de la o subpopulaţie la alta. De exemplu, gradul în care o trăsătură psihică este relevantă pentru subiecţi acţionează ca variabilă moderatoare dacă trăsătura trebuie să prezică un anumit comportament.

Variabilele mediatoare se folosesc atunci când există o relaţie strânsă între variabila independentă şi cea dependentă. Uneori, în spatele unei variabile moderatoare se află o variabilă mediatoare.

De exemplu: O cercetare privind anxietatea faţă de calculator a studenţilor psihologi (Pitariu, Albu, 1995) a evidenţiat faptul că între extraversie şi convingerea studenţilor că nu pot învăţa să lucreze la calculator (unul dintre factorii anxietăţii faţă de calculator) intervine, ca variabilă

mediatoare, modul în care subiecţii îşi apreciază cunoştinţele de informatică (nota pe care şi-o acordă). Variabila mediatoare se corelează pozitiv cu extraversia şi negativ cu factorul anxietăţii faţă de calculator.

Pe de altă parte, autoaprecierea subiecţilor se corelează liniar pozitiv cu nivelul real al cunoştinţelor de informatică ale acestora, care este strâns legat de profilul liceului absolvit (liceu de matematică-fizică, de filologie, de arte etc). Dacă cercetarea s-ar fi restrâns la mai puţine variabile, neluând în considerare nivelul cunoştinţelor de informatică autoapreciat sau apreciat obiectiv, s-ar fi constatat că profilul liceului absolvit joacă rol de variabilă moderatoare între extraversie şi anxietatea faţă de calculator.

In unele situaţii, o variabilă poate acţiona atât ca moderator, cât şi ca mediator. De exemplu, autoeficienţa (care a fost definită de A. Bandura ca „aprecierea făcută de individ asupra capacităţii proprii de a organiza şi executa acţiunile necesare pentru a atinge o anumită performanţă") funcţionează ca:

• mediator între situaţia de lucru apreciată din punctul de vedere al complexităţii lucrului şi al controlului deţinut de subiect şi iniţiativa personală; aceasta înseamnă că atât controlul pe care îl are subiectul asupra muncii, cât şi complexitatea lucrului au un efect direct asupra autoeficienţei sale şi aceasta are efect direct asupra iniţiativei personale;

• moderator în relaţia dintre situaţia de lucru şi iniţiativa personală; relaţia dintre controlul deţinut de subiect şi complexitatea lucrului, pe de o parte, şi iniţiativa personală, pe de altă parte, este dependentă de nivelul de eficienţă perceput de persoană (Speier, Frese, 1997).

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,

Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition)*, Simon&Schuster, Upper Saddle River, New Jersey,

Kline, P., 2000- *The handbook of psychological testing (Second Edtition)*,Routledge, London and New York,

Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului*, Ed. Fundatiei Romania de Maine, Bucuresti,

Mitrofan Niculae- Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare,* Editura Polirom, Iaşi

Urbina Susana, *2004- Essential of psychological testing*, John Wiley & Sons, Inc. Hoboken, New Jersey

6. CARACTERISTICI METROLOGICE ALE TESTULUI: FIDELITATEA

6.1. Definiţia fidelităţii
6.2. Rolul coeficientului de fidelitate şi eroarea standard a măsurării
6.3. Metode pentru evaluarea fidelităţii testelor
6.4. Observaţii asupra metodelor folosite pentru estimarea fidelităţii
6.5. Factorii care intervin în estimarea coeficienţilor de fidelitate

6.6. Interpretarea coeficienţilor de fidelitate

6.1. DEFINIŢIA FIDELITĂŢII

Termenul fidelitate- în limbajul curent se asociază cu încrederea. În ştiinţele sociale şi comportamentale, măsurătorile sunt mult mai predispuse la eroare datorită naturii intangibile a constructelor care sunt evaluate şi datorită faptului că o mulţime de factori dificil de controlat pot afecta comportamentele. Scorurile testelor psihologice sunt în special susceptibile să fie influenţate de o mulţime de surse: persoana examinată, examinatorul, contextul examinării – toate acestea pot determina variaţii străine sau opuse scopului testării.

1. 1954 şi 1966, în *Standards For Educational And Psychological Tests*- „Fidelitatea reprezintă precizia, consistenţa şi stabilitatea măsurării realizate de test”
2. 1974- în aceleaşi standarde „fidelitatea este definită prin gradul în care rezultatele testării pot fi atribuite surselor sistematice de variaţie”
3. 1985- în aceleaşi standarde „fidelitatea este gradul în care scorurile testului sunt consistente sau repetabile, adică gradul în care ele nu sunt afectate de erorile de măsură”
4. 2004- Susana Urbina- „În contextul testării erorile de măsură pot fi definite ca fiind orice fluctuaţie a scorurilor care rezultă din factori ai măsurării, irelevanţi pentru ceea ce se măsoară. Definim atunci fidelitatea ca o calitatea a scorurilor testului care sugerează că sunt suficient de consistente şi libere de erori de măsură astfel încât să fie folositoare.”

Fidelitatea este o caracteristică a scorurilor testului , mai degrabă decât a testului în sine. Fidelitatea oricărei măsurători, psihologice în particular, nu este absolută sau imuabilă. De aceea posibilele surse de eroare şi măsura în care se pot manifesta trebuie luate în considerare, estimate, şi raportate de fiecare dată când testele sunt aplicate.

Un test dispune de dovezi ale fidelităţii măsurătorilor dacă:

- Este lipsit de erori de măsură
- Orice persoană obţine scoruri egale la test şi la o formă paralelă a testului, când erorile de măsură la cele două teste sunt independente

Criterii pentru fidelitatea tehnică:

- Dispersia rezultatelor să fie mare
- Distribuţia rezultatelor să fie normală
- Rezultatele aceluiaşi subiect să fie asemănătoare la retestare
- Pentru testele colective, fidelitatea să fie mai mare decât la cele individuale

6.2. ROLUL COEFICIENTULUI DE FIDELITATE ŞI EROAREA STANDARD A MĂSURĂRII (SEM)

Coeficienţii de fidelitate se utilizează la compararea testelor.Eroarea standard a măsurării – SEM- este utilizată la interpretarea scorurilor testului- adică la determinarea intervalelor de încredere pentru scorurile reale ale subiecţilor.
Formula de calcul:

$$\textbf{SEM=}\quad \boldsymbol{\sigma}\times\sqrt{1-rtt}$$

σ= abaterea standard a scorurilor testului(deviaţie standard- SD)
rtt= coeficientul de fidelitate pentru lotul testat

Se poate observa că valoarea SEM este cu atât mai mică cu cât coeficientul de fidelitate este mai mare. Spre exemplu dacă scorul la un test de inteligenţă pentru un subiect este de IQ=110, iar SEM este 5, alături de scorul obţinut trebuie să luăm în considerare şi variaţia posibilă a scorurilor, adică ± 3SEM, pentru încrederea maximă(99%)= ± 15. Este posibil ca scorul real al subiectului să fie 110, dar la fel de posibil să fie oricare altul în intervalul de încredere $[95 la 125]$. Dacă încrederea ste de 95% atunci scorul se paote afla în intervalul 110 ± 2 SEM, adică între 100 şi 120.

Sursele de variaţie a erorii şi tipul de fidelitate care pune în evidenţă măsimea erorii:
1. eşantionarea timpului- **fidelitatea test- retest** pune în evidenţă stabilitatea în timp a rezulatelor unei persoane
2. eşantionarea conţinutului-**fidelitatea prin corelaţia între formele alternative**(aplicare imediată) pune în evidenţă erorile de eşantionare a conţinutului
3. eşantionarea conţinutului şi a timpului- **fidelitatea prin corelaţia între formele alternative**(la interval mai mare de timp)
4. eşantionarea conţinutului şi eterogenitatea acestuia (inconsistenţa între itemi) - fidelitatea calculată prin **metoda înjumătăţirii**, sau metoda **Kuder-Richardson**, sau **Alfa- Cronbach**
5. diferenţe între evaluatori- **fidelitatea interevaluatori** pune în evidenţă erorile datorate acestor diferenţe

6.3. METODE PENTRU EVALUARE FIDELITĂŢII TESTELOR

Metode:
- Metoda test- retest
- Metoda analizei consistenţei interne
- Metoda formelor paralele
- Coeficienţi de fidelitate interevaluatori

Pentru un test, în funcţie de ceea ce măsoară şi de condiţiile în care a fost administrat, se calculează unul sau mai mulţi dintre următorii coeficienţi, prin care se aproximează valoarea coeficientului de fidelitate:

A. **coeficienţii de stabilitate**- metoda test-retest (stabilitate în timp)
B. **coeficienţi de echivalenţă**- metoda formelor paralele, prin corelarea scorurilor obţinute la două teste paralele, administrate aproape în acelaşi timp(arată cât de asemănătoare sunt cele două instrumente de măsură)
C. **coeficienţii consistenţei interne**- pe baza scorurilor obţinute un test administrat o singură dată- coeficientul α al lui Cronbach, coeficientul lui Guttman, şi coeficienţii Kuder-Richardson(exprimă concordanţa diferitelor părţi ale testului)
D. **coeficienţii de fidelitate interevaluatori**- se calculează pentru teste al căror scor este rezultatul aprecierii subiective făcute de evaluator(arată în ce măsură părerile unor evaluatori concordă între ele).

Metoda test-retest

Este utilă când scorurile reale ale testului măsoară caracteristici durabile, generale şi specifice ale persoanelor. Se consideră că factorii temporari nu modifică scorul real al persoanei, ci influenţează erorile de măsură.

Procedură:

1. administrarea testului unui grup
2. după un interval de timp(nu mai puţin de trei luni, şi nu mai mult de şase luni- P.Kline, 2000) se administrează acelaşi test, aceloraşi persoane, în aceleaşi condiţii
3. se calculează coeficientul de corelaţie liniară între scorurile obţinute la cele două testări

La testele de aptitudini şi de cunoştinţe corelaţia dintre scoruri nu măsoară fidelitatea, uneori, măsoară schimbarea nivelului cunoştinţelor sau aptitudinilor.

Când valoarea coeficientului de corelaţie liniară test-retest este apropiată de 1- atunci se poate să existe şi situaţia când testul nu este potrivit pentru grupul de subiecţi- este foarte greu sau foarte uşor, deci intrepretarea coeficientului trebuie însoţită de analiza formei histogramelor scorurilor observate la test şi retest.

Când se comunică valoarea coeficientului de stabilitate se mai comunică şi:

- Lungimea intervalului de timp dintre test şi retest
- Condiţiile în care s-au efectuat testările(surse posibile de eroare)
- Vârsta subiecţilor testaţi

Atunci când testul este utilizat pentru decizii, prin plasarea subiecţilor în categoria „admis” sau „respins”(scorul mai mare sau mai mic decât pragul de admitere- cut off score)- este considerat stabil dacă nu se produc modificări în clasarea subiecţilor.

Dezavantaj- necesită timp şi cheltuieli materiale- pentru administrarea de două ori.

Metoda analizei consistenţei interne

Consistenţa internă înseamnă că itemii testului se referă la acelaşi lucru. Pentru a afla consistenţa se pot calcula unul din trei coeficienţi(cel adecvat situaţiei):

- Alfa α - al lui Cronbach(1943)- este utilizabil când subiecţii răspund pe o scală cu mai multe trepte(ex. 3, 5 sau 7, pentru cinci trepte- ex. niciodată - foarte rar - uneori - foarte des - întotdeauna, unde niciodată înseamnă 1 , întotdeauna înseamnă 5)

$$r_{xx} \text{ (sau } \alpha) = \frac{n}{n-1}\left(1 - \frac{\Sigma(SDi)^2}{SD_x^2}\right)$$ în care:

SD_x^2 = suma varianţelor la scorurile totale ale testului
$(SDi)^2$ = suma varianţelor la scorurile fiecărui item al testului
n = numărul de itemi ai testului

- Lamda λ_3 - al lui Guttman(1968)
- r 20 sau r 21 ai lui Kuder-Richardson(1939)- se aplica atunci când răspunsurile sunt de tipul DA-NU, Adevărat-Fals(binare)

$$r_{tt} = \frac{n}{(n-1)} \times \frac{SD_t^2 - \Sigma pq}{SD_t^2}$$, în care.

r_{tt} = coeficientul de fidelitate al întregului test
n = numărul de itemi ai testului
SD_t = deviaţia standard a scorurilor totale ale testului
Σpq = suma produselor persoanelor care reuşesc şi a celor care nu reuşesc la fiecare item

Metoda este utilă atunci când testul măsoară o caracteristică unidimensională, şi arată cât de bine fiecare item măsoară aceeaşi caracteristică. Testul trebuie administrat fără limită de timp, astfel încât toţi subiecţii să termine testul. Pentru o consistenţa internă bună trebuie aleşi itemii cu dispersii mici ale scorurilor şi corelaţii mari cu scorul total al testului.
Dacă un test are n itemi paraleli, atunci coeficientul său de fidelitate se măsoară cu formula lui Spearman- Brown:

$$\mathbf{r} = \frac{n_{y1,y2}}{1-(n-1)\times r_{y1,y2}}$$

Se observă că fidelitatea testului depinde de numărul itemilor componenţi(în teoria clasică a psihometriei). Din această formulă se poate afla de câte ori trenuie lungit sau scurtat testul, prin itemi paralei, astfel încât fidelitatea să fie egală cu o valoare fixată.

Metoda formelor paralele

Constă în administrarea a două teste, care reprezintă instrumente de măsură paralele, la momente foarte apropiate unul de altul, şi determinarea coeficientului de corelaţie liniară între scorurile observate ale celor două forme. Valoarea calculată se numeşte **coeficient de echivalenţă**. Dacă cele două teste sunt paralele şi dacă pentru fiecare persoană erorile de măsură la cele două administrări sunt variabile aleatoare independente, atunci coeficientul de echivalenţă coincide cu coeficientul de fidelitate al fiecărui test.

Metoda formelor paralele este **potrivită în cazul** când se **urmăreşte** măsurarea unor caracteristici *generale* **ale persoanelor.** Asemenea caracteristici vor influenta în acelaşi mod scorurile ambelor teste.

Metoda formelor paralele prezintă unele **avantaje** faţă de metoda test- retest: cele două forme fiind diferite, rezultatele la cel de-al doilea test nu sunt influenţate nici de memorarea răspunsurilor date la primul test şi nici de aflarea răspunsurilor la întrebările Ia care nu s-a ştiut răspunde (în cazul testelor de aptitudini sau de cunoştinţe).

Dar, metoda formelor paralele are **inconveniente.** Dintre acestea amintim următoarele:

- sunt necesare două administrări separate, ceea ce consumă timp;
- intervine oboseala sau plictiseala persoanelor examinate, ceea ce influenţează scorurile la cel de al doilea test
- elaborarea unor forme paralele poate fi dificilă şi costisitoare

Metoda se mai foloseşte şi atunci când se doreşte obţinerea unei colecţii de teste paralele, de bună calitate.

Un caz particular îl constituie **metoda înjumătățirii(half-split).** Uneori este considerată ca metodă de analiză a consistenţei interne. Pentru că e dificil să se creeze forme paralele se procedează altfel:

- Se administrează testul

- Se împarte în două jumătăţi cât mai asemănătoare
- Se calculează coeficientul de corelaţie între cele două jumătăţi.

Dacă jumătăţile testului sunt instrumente de măsură paralele atunci valoarea calculată reprezintă coeficientul de fidelitate al fiecăreia dintre ele. Aplicând formula Spearman- Brown se calculează coeficientul de fidelitate pentru întreg testul. Metoda nu se aplică atunci când administrarea testului se face cu limită de timp. Această metodă rezolvă multe din inconvenientele metodei test-retest şi a formelor paralele, dar are şi dificultăţi- itemii slectaţi în cele două jumătăţi trebuie să fie echivalenţi, şi independenţi.

Coeficientul de fidelitate interevaluatori

Se utilizează în cazul testelor care nu au o cotare obiectivă- cum ar fi cele proiective sau cele de creativitate. Pentru a vedea în ce măsură scorurile sunt dependente de cel care a făcut cotarea se calculează coeficientul de corelaţie liniară între scorurile atribuite aceloraşi subiecţi de către evaluatori diferiţi. Unii autori recomandă pentru aprecierea acordului între evaluatori a coeficientului K al lui Cohen, care ia valori între +1 şi -1, şi indică în ce măsură concerdanţa evaluărilor este diferită de cea datorată întâmplării.

6 .4. OBSERVAŢII ASUPRA METODELOR FOLOSITE PENTRU ESTIMAREA FIDELITĂŢII

Alegerea metodei aplicate trebuie să ţină seama de ceea ce **măsoară testul, de scopul** pentru care el var fi folosit şî **de condiţiile concrete** în care se poate face testarea (există sau nu posibilitatea administrării repetate a testului, există sau nu forme paralele ale testului etc).

Când testul măsoară o caracteristică durabilă şi, deci, trebuie să fie puţin sensibil la situaţia de testare, sau când pe baza scorurilor testului urmează să se tragă anumite concluzii valabile pentru o perioadă de timp mai îndelungată, se va aplica metoda test - retest.

In situaţia când nu se poate asigura retestarea persoanelor după un interval de timp suficient de lung pentru ca subiecţii să nu îşi mai poată aminti răspunsurile date la test, dar există o formă paralelă a testului, se va utiliza metoda formelor paralele. In cazul particular, când forma paralelă a testului a fost administrată împreună cu testul, se va putea aplica metoda înjumătăţirii.

Dacă se ştie că testul se compune din itemi care măsoară aceeaşi caracteristică - mai mult sau mai puţin durabilă - pentru estimarea coeficientului de fidelitate, se va folosi o formulă de calcul al unui coeficient de consistenţă internă.

Pentru testele unde nu se poate realiza o cotare strict obiectivă, evaluatorul devine o componentă a procesului de măsurare realizat cu ajutorul testului. Alături de o altă metodă de estimare a fidelităţii se va calcula şi coeficientul de fidelitate interevaluatori.

In general, **valorile coeficienţilor de consistenţă internă sunt mai mari decât cele ale coeficienţilor de echivalenţă, care sunt,** de obicei, **mai mari decât coeficienţii de stabilitate** (Murphy, Davidshofer, 1987). Acest lucru se explică prin faptul că numărul factorilor care influenţează erorile ele măsură este mai mare în metoda test - retesl şi în metoda formelor paralele, decât în metoda analizei consistenţei interne. De exemplu, acumularea de informaţii între test şi retest va afecta numai valoarea coeficienţilor de stabilitate, nu şi pe aceea a coeficienţilor de echivalenţă sau a coeficienţilor consistenţei interne.

Dacă, însă, persoanele examinate ţin minte răspunsurile date la test şi le repetă la retest, se poate obţine un coeficient de stabilitate mare, chiar dacă testul are consistenţă internă scăzută.

6.5. FACTORI CARE INTERVIN ÎN ESTIMAREA COEFICIENŢILOR DE FIDELITATE

Experimentele efectuate pentru determinarea valorii unui coeficient de fidelitate **nu cuprind,** în general, **întreaga populaţie de persoane căreia îi este destinat testul,** ci doar un eşantion din această populaţie. Din acest motiv, **valoarea calculată este doar o aproximare**(o estimare) a coeficientului de fidelitate. Selecţia unui alt lot va conduce la obţinerea unei alte valori.

Pentru creşterea preciziei coeficientului de fidelitate, este necesar să *fie* îndeplinite următoarele **condiţii** (Traub, 1994):

A. **Eşantionul să fie cât mai voluminos.** Deşi forma exactă a repartiţiei coeficienţilor de fidelitate calculaţi pentru diverse eşantioane de acelaşi volum, extrase aleator din populaţie, nu este cunoscută, se obişnuieşte să se considere că abaterea standard a acestei repartiţii (denumită eroare standard) este invers proporţională cu rădăcina pătrată din volumul eşantioanelor. Prin urmare, **pentru a micşora de două ori eroarea standard este necesar să se mărească de patru ori volumul eşantionului.**

B. Eşantionul de subiecţi să fie reprezentativ pentru populaţia căreia îi este destinat testul. El trebuie să fie **la fel de eterogen** ca şi această populaţie, întrucât gradul de omogenitate al eşantionului are influenţă asupra mărimii coeficientului de fidelitate.

C. Măsurătorile efectuate cu ajutorul testului **să fie independente între ele.** Măsurătoarea realizată asupra unui subiect să nu fie influenţată şi nici să nu influenţeze măsurătorile efectuate asupra altor persoane. Dacă administrarea testului se face în grup, acest lucru înseamnă că persoanele examinate vor fi supravegheate astfel încât nici una să nu poată copia.

Atunci când calculul coeficientului de fidelitate impune ca fiecare persoană să răspundă de două ori la acelaşi test sau să răspundă la două forme paralele ale testului foarte asemănătoare între ele, cerinţa de independenţă a măsurătorilor se extinde, în sensul că **măsurătorile efectuate asupra aceleiaşi persoane trebuie să fie independente între ele.** Aceasta înseamnă că nu li se va oferi subiecţilor posibilitatea de a revedea sau de a-şi aminti răspunsurile date la prima administrare a testului.

Dacă se calculează un coeficient de consistenţă internă, se- va urmări ca **itemii testului să fie independenţi între ei.** Prin urmare, rezolvarea unuia nu va fi condiţionată de rezolvarea altuia şi nu vor exista itemi care să se refere la acelaşi aspect, astfel încât subiecţii să răspundă la ei în acelaşi mod (corect sau greşit).

D. Procedura de administrare a testului şi a retestului sau a celor două forme paralele ale testului să fie **identică** şi, în acelaşi timp, să **fie aceeaşi cu cea care se va utiliza în practică.** Astfel, de exemplu, dacă se prevede ca testul să fie administrat cu limită de timp, în 40 de minute, atunci acest interval de timp va fi respectat şi la administrarea testului pentru estimarea fidelităţii.

6.6. INTERPRETAREA COEFICIENŢILOR DE FIDELITATE

Când avem de ales dintr-o mulţime de teste având toate celelalte caracteristici egale, este de preferat testul care are cea mai mare fidelitate. Dar, pe de o parte, testul cel mai fidel poate fi cel mai scump sau mai dificil de administrat. Pe de altă parte, caracteristicile testelor sunt rareori egale,

Nu este posibil să se specifice un nivel minim care să fie impus coeficienţilor de fidelitate ai tuturor testelor. In practica obişnuită sunt utile următoarele **recomandări:**

1. Se vor cere **nivele înalte** de fidelitate testelor folosite pentru a se lua **decizii importante asupra persoanelor şi celor care împart indivizii în mai multe categorii pe baza unor diferenţe relativ mici între** ei. In selecţia profesională, de pildă, se vor folosi teste care au **coeficienţi de fidelitate mai mari decât 0.90 (J.** C. Nunnally, după Hammond, 1995).

2. Se vor accepta **nivele scăzute** de fidelitate atunci când testele se utilizează pentru a lua **decizii preliminare, nu finale,** şi în cazul testelor care **împart persoanele într-un număr mic de categorii,** pe baza unor diferenţe individuale mari.

3o Dacă un test serveşte la **compararea grupelor de persoane, coeficientul său de fidelitate poate fi mai mic, chiar cuprins între 0.6 şi 0.7. Dacă, însă testul se foloseşte pentru compararea persoanelor, una cu alta, coeficientul său de fidelitate trebuie să fie mai mare decât 0.85** (Aiken, 1994).

Fidelitatea chestionarelor de personalitate şi a celor de interese tinde să fie mai mică decât cea a testelor cognitive (de cunoştinţe, de inteligenţă, de aptitudini speciale, Aiken, 1994).

In general, se întâlnesc de obicei coeficienţi de fidelitate cuprinşi între **0.70** şi **0.98** (Guilford, 1965).

Testele cotate obiectiv care măsoară aptitudini cognitive pot furniza, în loturi eterogene de persoane, coeficienţi de fidelitate mai mari decât 0.8.

Testele cotate subiectiv care măsoară aptitudini, testele de cunoştinţe şi cele care evaluează performanţe tipice (de exemplu, chestionare de personalitate) au, adesea, coeficienţi de fidelitate mai mici decât 0.80 (Traub, 1994).

Testele standardizate de cunoştinţe au, în general, o fidelitate mare spre moderată.

Testele cu alegere multiplă pot avea o fidelitate mai scăzută. Coeficientul lor de fidelitate este apropiat de 0.75.

In practică se folosesc **adesea teste care sunt valide pentru populaţia şi situaţia în care sunt administrate, dar care nu au coeficienţi de fidelitate foarte mari.**

Tipul de fidelitate necesar a fi calculat(Urbina, S., 2004), pe scurt:

- Dacă scorarea testului implică judecăţi subiective atunci trebuie calculată fidelitatea interevaluatori, (sau testul trebuie să aibă un coeficient ridicat)
- Dacă un test va fi utilizat pentru a măsura schimbarea în timp, cum ar fi îmbunătăţirile în urma unei intervenţii terapeutice, o estimare a erorii datorate eşantionării timpului este esenţială
- Dacă există posibilitatea ca o persoană să fie retestată ulterior, pentru a confirma rezultatele anterioare, atunci este necesară fidelitatea prin metoda testelor paralele
- Dacă omogenitatea sau consistenţa internă a testului este necesară, atunci avem nevoie de coeficientul KR-20(Kuder- Richardson) sau Alpha Cronbach.

Alte consideraţii despre fidelitate atunci când selectăm un test(Urbina, S., 2004):

Pasul 1. Determinăm sursele potenţiale de eroare care pot afecta scorurile instrumentelor studiate

Pasul 2. Examinăm datele despre fidelitate ale acestor instrumente, inclusiv tipurile de eşantioane pe care aceste date au fost obţinute

Pasul 3. Evaluăm datele despre fidelitate în lumina celorlalte atribute ale testelor- cum ar fi datele despre norme şi validitate, constrângeri de timp şi costuri, etc.

Pasul4. Când toate celelalte sunt egale selectăm testul care promite că furnizeză cele mai de încredere scoruri pentru scopurile testării şi pentru populaţia supusă testării(simultan) .

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,

Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,

Kline, P., 2000- *The handbook of psychological testing (Second Edtition),*Routledge, London and New York,

Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului,* Ed. Fundatiei Romania de Maine, Bucuresti,

Mitrofan Niculae- Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare,* Editura Polirom, Iaşi

Urbina Susana, *2004- Essential of psychological testing,* John Wiley & Sons, Inc. Hoboken, New Jersey

7. CARACTERISTICI METROLOGICE ALE TESTULUI: VALIDITATEA

7.1. DEFINIŢIA VALIDITĂŢII ŞI PROBLEME ACTUALE LEGATE DE VALIDITATE

Ca rezultat al evoluţiei psihometriei concepţia ştiinţifică despre validitate a fost de mai multe ori redefinită. La începuturile psihometriei validitatea era considerată o caracteristică a testului, sau după primii ani de avânt, era trecută pe planul secund. Primele teste psihologice ajunseseră să fie văzute, oarecum într-un mod naiv, ca instrumente ştiinţifice care măsurau un lung şir de abilităţi mentale şi trăsături de personalitate, cam în acelaşi fel în care psihofizicienii măsurau răspunsurile auditive, vizuale, sau alte răspunsuri senzoriale şi perceptive la stimuli de tip sunete, lumină şi culori de diferite tipuri şi intensităţi.

Mai apoi, succesul testelor Stanford-Binet şi Army Alpha în sprijinirea deciziilor practice despre indivizi, în educaţie şi angajări, a condus la o rapidă proliferare a testelor în primele două decenii ale secolului XX. Marea diversitate a aplicaţiilor pentru care aceste instrumente erau utilizate, foarte curând, a depăşit ca notorietate raţiunile ştiinţifice şi teoretice pentru ele, care erau disponibile la acea vreme. În scurt timp, multe teste psihologice au fost dezvoltate şi utilizate fără a beneficia de teoria psihometrică, principiile etice, şi ghidurile practice care au început să se acumuleze în ultimii zeci de ani ai secolului.(Urbina, S. 2004).

Testele psihologice există pentru a ne ajuta să facem inferenţe despre persoane şi comportamentul lor. Principalul pericol potenţial la adresa unui test este lipsa de reprezentativitate a constructului testului şi variaţia sa, precum şi lipsa de corectitudine a inferenţelor care se fac despre subiect pe baza scorurilor. Validitatea- de departe cea mai importantă problemă referitoare la scorurile testelor şi utilizările lor- depinde de dovezile pe care le putem aduce pentru a sprijini orice inferenţă care se cere făcută pe baza rezultatelor la test. Primordialitatea consideraţiilor de spre validitate este recunoscută în actualele standarde americane despre testare -Testing Standards(1999. AERA, APA, NCME)- prin plasarea sa ca subiect chiar la începutul acestora, în primul capitol.

Astfel, în **Testing Standards(**1999) este definită ca : **"gradul în care toate dovezile acumulate sprijină interpretarea dorită a scorurilor testului în raport cu scopul propus"**. În această definiţie sunt, implicit, trei idei interconectate care reflectă viziunea actuală a psihometriei asupra acestui concept central şi multifaţetat:

1. **Validitatea scorurilor la teste provine din toate dovezile acumulate** pentru a sprijini interpretarea şi utilizările lor. Astfel validitatea este întotdeauna o problemă de grad mai degrabă decât una de genul- "totul sau nimic". Validarea – procesul prin care dovezile despre validitate sunt adunate- începe cu o afirmaţie explicită a dezvoltatorului testului despre cadrul conceptual şi raţiunea testului, dar este prin natura sa fără limită deoarece include toate informaţiile care se adaugă în timp la înţelegerea rezultatelor testării.
2. Pe măsură ce se acumulează înţelegere teoretică şi dovezi empirice pentru interpretările scorurilor testului, **validitatea inferenţelor(a ipotezelor) făcute pe baza scorurilor pentru diverse scopuri propuse poate fi îmbunătăţită sau din contră- diminuată**. Un corolar al acestei noţiuni explicit afirmat în standardele de testare(1999), este că: "**validarea este responsabilitatea comună** atât a dezvoltatorului testului(care furnizează dovezi şi explicaţii pentru scopul de utilizare a testului) cât şi a utilizatorului testului(care evaluează dovezile disponibile în contextul concret pentru care testul este destinat)".
3. Datorită multiplelor scopuri pentru care pot fi utilizate testele, bazele pentru interpretarea scorurilor testului pot fi derivate printr-o **varietate de metode.** Contribuţiile la dovezile despre validitate pot fi făcute prin orice tip de cercetare sistematică care sprijină sau adaugă ceva la semnificaţia lor, indiferent cine desfăşoară cercetarea sau când este făcută. Atât timp cât există dovezi ştiinţifice de încredere pentru o utilizare propusă a scorurilor testului, utilizatorii calificaţi de teste sunt liberi să folosească scorurile pentru scopurile lor, indiferent dacă acestea au fost prevăzute de dezvoltatori sau nu. Această afirmaţie ajută la explicarea **multiplelor faţete** ale cercetărilor de validare, tot aşa cum se întâmplă deseori să fie constatări redundante iar alteori contradictorii. De asemenea este o explicaţie pentru longevitatea anumitor instrumente, cum ar fi MMPI, scalele de inteligenţă Wechsler, pentru care o vastă literatură- rezultatul numeroaselor aplicaţii într-o mare varietate de contexte- a fost acumulată de-a lungul a multor ani de cercetare de bază şi aplicată.

Din aceste explicaţii reiese că validitatea nu este o calitate a testului ci mai degrabă **este o problemă de judecăţi care sunt direct relaţionate cu scorurile testului,** în măsura în care acestea sunt o bază pentru un anume scop într-un context dat. Astfel procesul de validare este similar cu testarea ipotezelor : acesta subsumează semnificaţia scorurilor, fidelitatea scorurilor testului, la fel ca şi

modalităţile în care aplicaţiile informaţiillor testului la cercetările psihologice şi practice pot fi justificate.
Se poate afirma în acest moment că a existat o evoluţie a viziunii despre validitate de la una clasică la una modernă.
Validitatea testelor are patru faţete(Standardele APA, 1954), care au devenit patru stategii de validare a inferenţelor făcute pe baza scorurilor la test(Standardele APA, 1985,1999):

1. validitatea conceptuală
2. validitatea de conţinut
3. validitatea predictivă
4. validiatea concurentă

Principala evoluţie priveşte punerea în legătură directă validităţii cu scorurile obţinute la teste.
Validitatea unui test, arată Silva, se referă la cât de potrivite sunt interpretările descriptive, explicative sau predictive care se dau scorurilor sale (Silva, 1993). Într-o accepţiune actuală, conceptului de validitate i s-a dat o semnificaţie nouă. prin aceasta înţelegându-se „măsura în care o anumită inferenţă făcută plecând de la test are înţeles şi este potrivită" (APA *Standards* 1985).

Definiţie: Validarea unui test este procesul prin care se investighează gradul de validitate a **interpretării** propuse de acesta (APA,1985). Iată, deci, că validarea nu se mai referă atât la instrumentul în sine, cât la interpretarea datelor furnizate de el. Dar, pentru o interpretare corectă a scorurilor unui test, este necesar să fie îndeplinite două condiţii (Albu, 1995):
• să se ştie exact ce reprezintă scorurile testului;
• măsurările făcute de test să fie corecte.

C. L. Lawshe propune înlocuirea denumirii de „tip de validitate" cu cea de **„tip de analiză a validităţii".** Prin „analiza validităţii"", el înţelege o procedură, un proces sau o strategie prin care se colectează sau se generează date în vederea determinării extinderii ori a robusteţii inferenţelor care pot fi făcute pe baza unei mulţimi de scoruri de test (1986).
În linii generale, validitatea exprima gradul în care un test măsoară ceea ce îşi propune să măsoare. La această calitate se mai poate adăuga şi o alta, dacă testul poate fi utilizat în luarea unor decizii corecte. Cu alte cuvinte, daca noi cunoaştem performanţele unei persoane la un test (**predictor**), cât de precis vom putea estima ce performanţe profesionale va obţine?
Rezultatul este cunoscut ca şi **coeficient de validitate**. Un test poate avea mai mulţi coeficienţi de validitate, în funcţie de numărul de dimensiuni profesionale (calitatea muncii, disciplina, categoria profesionala etc.) care corelează cu el.
Pentru a sublinia ideea că este greşit să se spună “testul are validitate de un anumit tip” încă din 1985, în Standardele APA se modifică terminologia utilizată în legătură cu validitatea, înlocuind denumirile cunoscute ale tipurilor de validitate cu formulări mai adecvate ştiinţific: **“ dovezi ale validităţii relative la – conţinut, construct, criteriu”.** Coeficientul de validitate nu permite astfel să se afirme că testul are validitate relativă la criteriu ci constituie doar una dintre dovezile validităţii relative la criteriu. Din această perspectivă în cadrul acestei lucrări- utilizăm termenul de validitate predictivă, sau de construct(spre exemplu), ca abreviere convenţională **“a dovezilor validităţii relative la criteriu(predictivă), sau relative la construct.”**
Este absolut necesar ca atunci când se vorbeşte despre dovezile validităţii testului să se indice despre ce fel de validitate se vorbeşte(la criteriu- predictivă/ concurentă, sau descriptivă- de conţinut / conceptuală).Dacă în cadrul unei populaţii există subgrupuri cu caracteristici diferite, este posibil ca între acestea să apară deosebiri şi în privinţa validităţii predictive a testelor. Spre exemplu, un test poate fi un bun predictor al performanţei numai pentru bărbaţi, nu şi pentru femei, sau numai pentru persoane care au studii superioare. O variabilă care modifică validitatea predictivă, cum sunt variabilele din exemplele anterioare(sexul şi studiile), se numeşte variabilă moderatoare.

Încă din 1974 şi până în prezent, în majoritatea lucrărilor de specialitate, se vorbeşte de trei faţete ale validităţii:

- Relativă la construct
- Relativă la conţinut
- Relativă la criteriu(considerându-se că validitatea predictivă şi cea concurentă de referă la acelaşi lucru- diferenţa fiind doar de strategie)

În paralel s-au facut nişte alegeri metodologice ca urmare a efortului de clarificare conceptuală a psihometriei. Una dintre temele consistente ale acestui efort a fost integrarea majorităţii formelor de **dovezi ale validităţii**, ca **aspecte ale validităţii de construct** (Guion, 1991; Messick, 1980, 1988,1989; Tenopyr, 1986), ceea ce a dus la necesitatea reexaminării termenului de construct(recomandăm în acest sens lucrarea Urbinei S. 2004).

Susana Urbina(2004) subliniază **funcţia integrativă a constructului în validarea testelor,** şi citează analiza făcută de Embretson în 1983. Constructul este utilizat în două ipostaze: ca trăsătură, proces, caracteristică a cărei prezenţă sau absenţă rezultatele testului o certifică şi ca inferenţă ce poate fi făcută pe baza scorurilor testului.

Embretson (1983)(apud Urbina, S.2004) a propus o separare între cele două aspecte ale cercetărilor de validare relativă la construct

- O parte numită reprezentarea constuctului(construct representation)
- O altă parte numită- anvergura, sau raza de acţiune nomotetică(nomotetic span)

Potrivit autoarei menţionate cercetarea relativă la reprezentarea constructului este focalizată pe identificarea mecanismelor teoretice care subordonează performanţa în sarcină. Din perspectivă informaţională scopul reprezentării constructului este o descompunere a sarcinii. Aceasta poate fi aplicată unei varietăţi de sarcini cognitive, incluzând inferenţele interpersonale şi judecăţile sociale, şi presupune o examinare a răspunsurilor din punctul de vedere al proceselor, strategiilor, şi cunoştinţelor implicate în performanţă, deci este din punctul de vedere al executării sarcinii.

Raza de acţiune nomotetică se ocupă cu reţeaua de relaţii a unui test cu alte măsurători. Se referă la tăria, frecvenţa şi patternuri de relaţii semnificative între scorurile la teste şi alte măsuri ale aceloraşi sau altor trăsături, între scorurile la test şi măsurile criteriului etc.

Embretson (1983) descrie caracteristici suplimentare ale conceptelor de “reprezentare a constructului” şi “anvergură nomotetică” care ajută la clarificarea diferenţelor dintre cele două aspecte ale cercetărilor de validare.

Două sunt utile în contextul dovezilor referitoare la sursele dovezilor validităţii:

- Cercetările despre reprezentarea constructului sunt preocupate primordial cu identificarea diferenţelor în sarcinile testului, în timp ce tipul complementar de cercetări, legate de anvergura nomotetică sunt preocupate cu diferenţele dintre examinaţi. Descompunerea unui proces poate să nu aibe legătură cu variabilitatea sa în populaţia de referinţă- deci testul are o bună reprezentare a constructului dar nu discriminează. Iar deciziile asupra oamenilor se bazează pe variabilitate.
- Validarea aspectelor de reprezentare a constructului este independentă de dovezile care sprijină anvergura nomotetică a scorurilor, şi invers. Putem şti precis ce procese sunt implicate în performanţă dar dacă nu sunt corelaţii semnificative cu comportamente sau măsuri din afara testului atunci scorurile testului pot avea o utilizare limitată.La fel, este posibil să obţinem o puternică reţea de relaţii între scoruri şi alte măsuri, fărăr a avea o definire clară a constructului pe care scorurile îl reprezintă. Exemplul dat de autoare aici este cel al testelor de inteligenţă, care au o mare anvergură nomotetică(corelează într-o măsură mai mare sau mai mică cu o varietate de alte măsuri), dar încă au o bază teoretică neclară(reprezentarea conceptului- de aceea Boring, în 1923, afirma că inteligenţa este”ceea ce testele de inteligenţă măsoară”).

Schema conceptuală trasată de Embretson reţine noţiunea de **validare relativă la construct ca un mod unitar şi comprehensiv de a aborda ştiinţific integrarea tuturor dovezilor care privesc**

semnificaţia şi interpretarea scorurilor la teste, şi, în acelaşi timp furnizează o bază pentru a distinge între:

- Sursele de dovezi pentru validitatea scorurilor -care au legătură primordial cu a şti ce anume măsurăm(reprezentarea constructului)
- Sursele de dovezi pentru validitatea scorurilor- care au legătură cu inferenţele pe care le putem face pe baza a ceea ce măsurăm(anvergura nomotetică).

Aceste surse ale dovezilor de validitate pot fi, deseori sunt, interrelaţionate şi ambele implică elemente teoretice şi observabile, ca şi modele sau postulate privind interrelaţia dintre elemente.

Includem pentru simplitatea ei o clasificare făcută anterior acestor dezbateri de către Guion (1976), care distinge numai două tipuri de validităţi:

- **Validităţi în raport cu criteriul**(**concurentă** şi **predictivă**) prin care se stabileşte relaţia dintre scorurile la test şi cele de la criteriu (profesionale, şcolare etc.).
- **Validităţi descriptive** (**de conţinut** şi **conceptuală-de construct**) care evaluează sensul intrinsec al scorurilor la test.

După cum se vede în accepţiunea mai veche validitatea relativă la construct se referea doar la domeniul de conţinut la testului- reprezentarea conceptului, cum spune Embretson.

7.2 FAŢETELE VALIDITĂŢII: VALIDITATEA RELATIVĂ LA CRITERIU ŞI VALIDITĂŢI DESCRIPTIVE- VALIDITATEA RELATIVĂ LA CONSTRUCT ŞI VALIDITATEA RELATIVĂ LA CONŢINUT

În această lucrare am preferat prezentarea informaţiilor despre validitate în maniera clasică, viziunea celor trei tipuri mari de dovezi ale validităţii scorurilor. Deşi în practică validarea relativă la construct precedă relativă la conţinut şi criteriu, din motive legate de uşurinţa înţelegerii am preferat prezentarea lor în ordine inversă.

Validitatea relativă la criteriu(validarea relativă la criteriu)
Este gradul de încredere în inferenţele pe care le facem privitor la o performanţă a subiectului, sau la un comportament prezis, pe baza scorurilor obţinute la test.
Cu alte cuvinte, daca noi cunoaştem performanţele unei persoane la un test (**predictor**), cât de precis vom putea estima ce performanţe profesionale va obţine (denumită **criteriu**)?
Aprecierea validităţii relative la criteriu se bazează pe valorile coeficientului de validitate al testului, ca şi coeficient de corelaţie între scorurile testului şi măsurătorile criteriului(ex. performanţa în muncă).
Se poate ca variabila criteriu să fie un alt test psihologic, când se doreşte validarea unui dintre ele- şi atunci vorbim de validare concurentă.
O aceeaşi variabilă poate juca rolul de predictor într-un caz(spre exemplu: performanţa şcolară – predictor pentru performanţa profesională), şi rolul de criteriu în alt caz(performanţa şcolară prezisă de rezulatele unui test de inteligenţă).
Procedee utilizate pentru validare:

- Validarea predictivă
- Validarea concurentă
- Validarea prin grupele contrastante

Validarea predictivă- este procedeul prin care se caută dovezi privind eficienţa predicţiei comportamentului unei persoane într-o situaţie dată, pe baza scorurilor testului(eficienţa măsurată după un timp mai lung sau mai scurt, dar nu în acelaşi timp sau imediat după testare). Spre exemplu în

selecţia profesională- pentru estimarea validităţii predictive se calculează coeficientul de corelaţie liniară- r-între scorurile testului şi măsurătorile performanţei profesionale după un anumit timp, suficient ca persoanele să înveţe profesia. Atunci când un test nu are dovezi privind validitatea predictivă deşi se poate aplica în selecţie rezultatul său nu poate fi luat în considerare, ci numai utilizat pentru calculul coeficientului de validitate.

Validarea concurentă- este procedeul prin care se caută dovezi privind eficienţa predicţiei comportamentului unei persoane, utilizând în testare persoane cu vechime în munca respectivă, pentru care deţinem evaluări ale performanţelor în momentul testării. Se numeşte concurentă deoarece se face în acelaşi timp cu testarea, odată ce dipunem de date privind performanţa(este procesul invers-ştim predicţia posibilă şi observăm dacă scorurile testului ar fi confirmat-o). Se calculează tot coeficientul de corelaţie liniară între scorurile la test şi performanţele cunoscute în momentul testării.

Deoarece validarea concurentă se face pe un eşantion de populaţie deja selecţionat(personal deja angajat, sau studenţi deja admişi în facultate) apare un fenomen care influenţează coeficientul de validitate- fenomenul de restrângere a mulţimii, ceea ce presupune analiza atentă a semnificaţiei rezultatelor obţinute la test.

Fenomenul restrângerii mulţimii

Este un fenomen care trebuie luat în calcul în validarea concurentă, deoarece lotul pe care se validează testul nu este un lot care este supus selecţiei(din care în mod firesc unii ar fi fost eliminaţi), ci este un lot deja selecţionat, sau auto-selecţionat. Spre exemplu într-un loc de muncă, pe parcursul a trei ani, au rămas numai cei satisfăcuţi, motivaţi şi care au performanţe cel puţin satisfăcătoare. O parte din cei cu performanţe foarte bune şi potenţial de creştere au fost deja promovaţi. Dacă dorim să validăm un test pe care să-l utilizăm pentru predicţie, ne lipsesc rezultatele celor slabi şi ale celor promovaţi. Astfel coeficientul de corelaţie calculat are o altă valoare decât ar fi avut şi cu cei menţionaţi. În aceste condiţii se poate să obţinem coeficienţi de corelaţie mai mici decât în realitate.

Unii dintre specialişti în teoria testelor recomandă ca testele ce sunt utilizate în predicţie să fie validate predictiv iar cele utilzate în diagnoză să fie validate concurent.

Coeficientul de validitate al unui test este egal cu coeficientul de corelaţie liniară între predictor şi criteriu(în general).

În literatura de specialitate există şi recomandarea de a calcula coeficientul de validitate folosind alţi indici în locul coeficientului de corelaţie liniară. Anastasi(1976) propune ca atunci când testul şi criteriul sunt variabile continue să rămână coeficientul de corelaţie liniară, iar în celelalte cazuri să se folosească alte tipuri de coeficienţi. Cronbach(1966) permite calcularea cu ajutorul coeficientului de corelaţie a rangurilor; dacă regresia dintre test şi criteriu este o funcţie monotonă crescătoare, coeficientul de corelaţie a rangurilor are o valoare apropiată de 1, deci testul este valid în sensul lui Cronbach(în majoritatea cazurilor de selecţie profesională este vorba de comparaţia a două clasamente, criteriul fiind rareori pe o scală de interval). Unii autori(Hammond, 1995) recomandă ca validarea unui test relativă la criteriu să utilizeze mai multe variabile criteriu, dintre care unele nu au legătură cu ceea ce prezice sau măsoară testul. Este de dorit ca testul să nu coreleze cu ele.

Validarea prin compararea grupelor contrastante(putem spune că e o formă a strategiei de validare concurentă)

Pe baza valorilor criteriului(ex. performanţa profesională) se formează două grupe – una cu persoane ce au rezultate foarte bune, cealaltă cu persoane ce au rezultate foarte slabe. Dacă în urma **comparaţiei statistice a mediilor** rezultatelor obţinute la test se constată că acestea diferă semnificativ- se apreciază că testul este valid relativ la criteriu- şi are putere predictivă. Această metodă se aplică atunci când variabila – criteriu este calitativă, deci nu se poate calcula coeficientul de corelaţie liniară şi atunci când criteriul este compus din mai multe variabile- cantitative şi/ sau calitative. Este echivalent dacă afirmăm că pe baza rezultatelor testului putem diferenţia între cei foarte slabi şi cei foarte buni.

Problema criteriilor

Este una dintre problemele esenţiale ale validării. Printre cele mai cunoscute criterii ce sunt utilizate în practica psihodiagnostică sunt următoarele (Anastasi, A., 1988, 1996, apud Mitrofan, N. 2000):
1) Achiziţiile academice, evidenţiate prin note, medii, evidenţieri speciale, rezultate la concursuri, promovări, graduări, recompense, burse etc., sunt foarte frecvent folosite pentru validarea testelor de inteligenţă. De aceea, mai mulţi autori au considerat aceste teste ca fiind modalităţi de măsurare a aptitudinilor şcolare. De asemenea, achiziţiile academice sunt folosite şi pentru validarea testelor de măsurare a personalităţii şi a testelor pentru măsurarea aptitudinilor multiple;
2) Performanţa în cadrul unui domeniu de formare specializat este folosită în calitate de criteriu pentru validarea unor teste folosite pentru măsurarea aptitudinilor speciale. Pot fi date ca exemple: - achiziţiile finale în cadrul cursurilor de formare profesională specială sunt folosite pentru validarea testelor ce măsoară aptitudinile tehnice; - performanţa în anumite ramuri artistice pot fi folosite în vederea validării testelor pentru măsurarea aptitudinilor artistice;
3) Performanţa muncii (job performance) apare în calitate de criteriu folosit pentru validarea, în primul rând, a testelor pentru aptitudini speciale. Totodată, ea este utilizată, dar în mai mică măsură, pentru validarea testelor de inteligenţă generală şi a testelor de personalitate;
4) Diagnoza psihiatrică este folosită în calitate de criteriu pentru evidenţierea validităţii testelor de personalitate, dacă este bazată pe observarea prelungită şi detaliată a cazului;
5) Corelaţiile dintre un test nou şi testele disponibile anterioare; de exemplu, validarea testelor de grup folosind bateria Stanford-Binet;
6) Metoda grupurilor contrastante, ce implică un criteriu compozit, şi despre care am vorbit anterior. De exemplu, validitatea unor teste de aptitudini muzicale sau a unor teste de aptitudini tehnice pot fi verificate prin compararea scorurilor obţinute de studenţii admişi în instituţiile de artă sau de inginerie cu scorurile obţinute de cei respinşi la examenele de selecţie.

Validitatea testului relativă la constructul măsurat(validarea relativă la construct)

Două observaţii importante făcute de *Standardele* APA din 1985, şi menţinute şi în cele din 1999 relativ la „tipurile" de validitate:
• prin „tipuri" de validitate nu se înţeleg „categorii disjuncte";
• nu se poate afirma că un tip de validitate este mai potrivit decât altul pentru o anumită utilizare a testelor sau pentru o categorie specifică de inferenţe bazate pe scoruri.

Prin intermediul testelor psihologice se măsoară atribute abstracte ca inteligenţa, motivaţia, agresivitatea, depresia etc, denumite constructe. Asemenea variabile sunt construite pornind de la fapte observabile, printr-un şir de raţionamente bazate pe ipoteze şi deducţii. Din acest motiv, **validitatea relativă la construct este adesea denumită şi „validitate ipotetico-deductivă"** (Bacher, 1981). Toate constructele sunt însă conectate la realitate, reprezintă aspecte sau evenimente ale realităţi observabile. Constructele psihologice sunt evidenţiate, direct sau indirect, de comportament sau în urma conducerii unor experimente dedicate măsurării lor.
Deoarece termenii construct şi concept sau noţiune sunt adesea utilizaţi ca sinonimi, validitatea relativă la construct a testului este denumită şi **„validitate conceptuală"**. Intre cei doi termeni există însă o deosebire, evidenţiată de L. J. Cronhach :
Constructul este „o categorie creată intenţionat pentru a organiza experienţa în enunţuri generale care se prezintă sub formă de legi" (Si1va, 1993). Un construct este, deci, o noţiune elaborată şi utilizată într-un cadru ştiinţific (teoretic, metodologic sau aplicativ).
Validarea testului relativă la constructul măsurat se ocupă de calităţile psihice care contribuie la formarea scorurilor acestuia şi urmăreşte înţelegerea dimensiunilor evaluate de test. Privind constructul ca baza interpretativă a răspunsurilor subiecţilor, acest tip de validare are şi scopul de a identifica şi analiza procesele psihice declanşate sau detectate de test (Silva, 1993).
L. J. Cronbach, reluând ideea formulată de *Standardele* APA din 1954 conform căreia prin validarea testului relativă la constructul măsurat se validează atât testul, cât şi ipotezele aflate la baza sa, scrie: **„Validitatea testului şi validitatea constructului strict inseparabile**. Când un test nou este elaborat

pentru măsurarea unui construct cunoscut, riscul ca validitatea să fie nesatislacătoare este mai mare pentru test decât pentru construct. Se poate constata şi necesitatea revizuirii constructului. Un exemplu în acest sens se constituie abandonarea ideii tradiţionale asupra debilităţii mintale" (Silva, 1993).
Numai înţelegând constructul măsurat de test se pot formula ipoteze asupra predicţiilor pe care le face testul şi asupra relevanţei şi a reprezentativităţii conţinutului său.
Deci, în practică, **validarea constructului trebuie să preceadă validarea conţinutului testului şi validarea relativă la criteriu.**

Procedură: Pentru a putea verifica dacă un test măsoară bine un anumit construct. este necesar să se realizeze **o descriere a constructului în termeni comportamentali concreţi.** Operaţia se numeşte **„explicarea constructului"** şi constă din trei paşi (Murphy şi Davidshofer, 1991):

1. Se identifică acele comportamente care au legătură cu constructul măsurat de test.
2. Se identifică alte constructe şi pentru fiecare se decide dacă are sau nu legătură cu constructul măsurat de test.
3. Se alcătuieşte câte o listă de comportamente prin care se manifestă aceste constructe. Pentru fiecare din ele, pe baza relaţiilor dintre constructe, se decide dacă are sau nu legătură cu constructul măsurat de test.

Exemplu: Murpy si Davidshofer (1991) exemplifică procesul de descriere a constructului de „agresivitate la elevi" prin următorul exemplu:

Construct	• ***Comportament***
1. Identificarea comportamentelor referitoare la agresivitate	
agresivitate	atacă alţi elevi este capul răutăţilor domină în jocuri
2. Se identifică alte constructe şi se decide dacă au sau nu legătură cuacest comportament	
trebuinţă de putere agresivitate onestitate	atacă alţi elevi este capul răutăţilor domină în jocuri
3. Se identifică comportamentele aferente fiecărui construct şi sedetermină relaţia cu constructul care trebuie măsurat	
trebuinţă de putere agresivitate onestitate	ia decizii în grup atacă alţi elevi este capul răutăţilor domină în jocuri se abţine de la certuri spune adevărul profesorului

Din analiza modelului de mai sus rezultă un sistem de relaţii care înglobează un set de constructe şi comportamente, denumit de L. J. Cronbach, „reţea nomologică" („nomologică" în sensul că ea are la bază declaraţii exprimate sub formă de legi). Unii autori au folosit termenul de „stohastic", susţinând că relaţiile dintre componentele reţelei sunt mai degrabă probabilistice decât de tip lege.
Ideile referitoare la caracterul structural al constructelor şi la existenţa reţelelor nomologice au condus la descompunerea operaţiei de validare a testului relativă la constructul măsurat în trei componente:
• **validarea materialului testului** *(substantive validity)*, care necesită specificarea constructului

măsurat de test şi se confundă uneori cu validarea conţinutului testului;
• **validarea structurii constructului** aflat la baza testului *(structural validity),* care identifică elementele componente ale constructului şi investighează legăturile dintre ele;
• **validarea externă** *(external validity),* care urmăreşte să determine relaţiile constructului cu alte măsurători (constructe sau variabile observate).

Validarea externă se referă la calitatea testului de a avea relaţii corecte cu alte măsuri psihologice; ea este denumită uneori **„validitate convergentă şi discriminantă"** (Murphy şi Davidshofer, 1991):

a. Un test are **validitate convergentă** dacă măsoară ceea ce evaluează şi alte teste sau variabile care se referă la acelaşi construct, deci scorurile sale variază (liniar sau nu) în acelaşi sens cu rezultatele respectivelor măsurători.

b. Un test are **validitate discriminantă** dacă evaluează altceva decât diverse teste sau variabile despre care se ştie că se referă la constructe ce nu au legătură cu constructul măsurat de test. Aceasta înseamnă că scorurile testului nu sunt în relaţie funcţională monotonă (liniară sau nu) cu rezultatele acelor măsurători.

Procedură: Verificarea **validităţii convergente şi discriminante** se realizează conform următorului algoritm:

• Se identifică acele comportamente care sunt legate de constructul măsurat de test.
• Se determină constructele care au legături cu constructul măsurat ele test şi se pun în evidenţă acele comportamente legate de ele care pot fi conectate cu constructul măsurat de test.
• Pentru toate comportamentele reţinute se apreciază ce fel de corelaţii ar putea exista între măsurile lor şi o măsură bună a construcţii lui care ne interesează (corelaţii pozitive, negative, mari, mici sau nule).
• Se determină coeficienţii de corelaţie între scorurile la testul studiat şi rezultatele unor măsurători (eventual, scoruri de teste) efectuate asupra comportamentelor reţinute.
• Se compară coeficienţii de corelaţie obţinuţi cu valorile aşteptate.

Procedură: O metodă utilizată pentru **verificarea validităţii convergente şi discriminante a testelor** este aceea propusă de D. T. Campbell şi I). W. Fiske, **bazată pe matricea multitrăsătură-multimetodă (***multitrait-multimethod,* Anastasi, 1976; Pitariu, 1994). Aceasta conţine valorile coeficienţilor de corelaţie liniară între diverse măsurători ale aceloraşi constructe, dar obţinute prin intermediul a două sau trei tehnici ori instrumente diferite.
Exemplu:În urma evaluării constructelor **a,b,c** , despre care se face presupunerea că nu au legături între ele, prin trei metode diferite „test *I*”, „test *II* şi „test III, s-a obţinut următoarea matrice de coeficienţi de *corelaţie liniară* (pentru a se putea urmări mai uşor datele, ea a fost scrisă sub forma unui tabel unde coeficienţi de corelaţie sunt la pragul p = 0,05):

Tabel privind matricea multitrăsătură- multimetodă, (Campbell şi Fiscke)

Metode / constructe	**Test I** **a b c**	**test II** **a b c**	**test III** **a b c**
test I **a** **b** **c**	 .81 .12 .86 .04 .32 .76		
test II **a** **b** **c**	**.46** .14 **.48** .08 .21 **.55**	.74 .02 .77 .10 .15 .82	

test III **a** **b** **c**	 **.50** .32 **.46** .28 .36 **.60**	 **.54** .37 **.65** .28 .38 **.61**	 .85 .38 .83 .36 .40 .85

Valorile notate pe diagonala principală sunt coeficienţii de fidelitate ai testelor, obţinuţi prin metoda formelor paralele. De exemplu, pentru testul **II,** coeficienţii de fidelitate sunt .74 (pentru constructul A), .77 (pentru constructul B) si .82 (pentru constructul C). Valorile de pe diagonalele celorlalte căsuţe sunt coeficienţii de corelaţie între măsuralorile realizate de cele trei teste pentru aceleaşi constructe. Ei servesc la verificarea **validităţii congruente.**

De exemplu, pentru **constructul a,** coeficienţii de corelaţie sunt:

- **.46 (testul I cu testul II);**
- **.50 (testul I cu testul III);**
- **.54 (testul II cu testul III).**

Pentru **constructul b:**

- **.48 pentru testul I cu II**
- **.46 pentru I cu III**
- **.65 pentru II cu III**

Pentru **constructul c:**

- **.55 pentru I cu II**
- **.60 pentru I cu III**
- **.61 pentru II cu III**

Toţi aceşti coeficienţi sunt semnificativi statistic la pragul p .05, ceea ce denotă **validitatea convergentă** a celor trei teste.

Celelalte valori din matrice reprezintă coeficienţii de corelaţie liniară între constructe diferite, măsurate prin acelaşi test sau prin teste diferite. Ei servesc la aprecierea **validităţii discriminante** a testelor.

Se observă că **testele I şi II discriminează între constructele A şi B** (coeficienţii de corelaţie aflaţi sub diagonala căsuţelor I - I, I - II şi II - II sunt nesemnificativi la pragul p .05), în timp ce testul III nu are validitate discriminantă (coeficienţii de corelaţie liniară între constructele A, B şi C, măsurate prin testul III - în căsuţa III - III - sau prin testul III şi un alt test - în căsuţele I—III şi II - III - sunt, în general, semnificativi la pragul p .05).

In exemplul prezentat, validitatea convergentă şi discriminantă a testelor **I, II şi III** a fost uşor de dedus. Dar, în practică, apar situaţii complexe, când mulţi coeficienţi de corelaţie liniară pentru un acelaşi construct sunt nesemnificativi şi mulţi coeficienţi de corelaţie liniară între constructe diferite sunt semnificativi statistic. În asemenea cazuri este greu de apreciat care dintre teste este lipsit de validitate.

Dacă testele măsoară constructe cu o structură complexă, pot să apară dificultăţi la interpretarea scorurilor. Din acest motiv, unii autori, ca M. L. Tenopyr, D. W. Fiske, J. C. Nunnally şi R. L. Durham, susţin că pentru a putea utiliza la interpretarea scorurilor unui test deducţii bazate pe felul în care este construit testul, trebuie ca testul să se refere la constructe simple, bine definite şi cu manifestări uşor de observat (Silva, 1993).

Procedură: O altă metodă de studiere a validităţii relative la construct implică **manipularea experimentală a constructului măsurat de test (Murphy, 1987).** De exemplu în cazul unui test destinat să măsoare anxietatea, se formează două grupe similare de subiecţi. În condiţii care nu generează anxietate cele două grupe vor realiza la test scoruri asemănătoare. Dacă, însă la administrarea unui grup i se spune ca va avea o surpriză neplăcută, iar celuilalt i se face instructajul obişnuit, este de aşteptat ca scorurile testului să diferenţieze semnificativ între grupe. Trebuie să se reţină că validitatea relativă la construct a unui test nu poate fi exprimată printr-o mulţime de coeficienţi. Ea se deduce prin acumularea rezultatelor cercetării relative la ipotezele propuse(Silva, 1993).

Validitatea unui test relativă la conţinutul său

Se referă mai ales la categoria testelor de achiziţie. Spunem că un test are validitate de conţinut, dacă el măsoară ceea ce şi-a propus să măsoare şi dacă elementele sale de conţinut (itemii) sunt expresia unui eşantion reprezentativ pentru un anumit univers de itemi sau univers de sarcini.

Definiţie: Acest tip de validitate este analizat la testele care se utilizează pentru a estima ”cum acţionează o persoană în universul de situaţii pe care testul intenţionează să-l reprezinte”.

Observaţii: În literatura psihologică există păreri diferite despre ceea ce se urmăreşte prin validarea testului relativă la conţinut. Anastasi(1976) consideră că acest tip de validare implică examinarea sistematică a conţinutului testului pentru a determina dacă el acoperă un eşantion reprezentativ din domeniul de comportamente pe care testul în cauză trebuie să-l măsoare. Dacă eşantionul este reprezentativ pentru populaţie, atunci se admite că testul are validitate relativă la conţinut(Gregory, 1992). Alţi autori înlocuiesc expresia „domeniu de comportamente” cu „mulţimea sarcinilor”, sau „mulţimea performanţelor”. M.M. Linchan susţine că validitatea relativă la conţinut determină cât de bine reprezintă condiţiile în care este observat (prin test) comportamentul persoanei toate mulţimile de condiţii. R.T. Lennon consideră că, validitatea relativă la conţinut se referă la răspunsurile subiectului şi nu la întrebările testului.

Pentru analizarea acestui tip de validitate trebuie luat în considerare nu numai conţinutul itemilor testului, ci şi procesul utilizat de subiect pentru a ajunge la formularea răspunsului(Silva, 1993). De exemplu, un test având drept domeniu de conţinut „cunoştinţele de ortografie”, format din itemi cu alegere multiplă, poate verifica dacă un individ recunoaşte cuvintele scrise corect din punct de vedere ortografic, dar nu şi dacă acesta ştie să scrie corect după dictare. Deci, nu este suficient să se spună că testul evaluează cunoştinţele de ortografie. În unele accepţiuni, validarea relativă la conţinut este o componentă a validării relative la construct. Ea se realizează după definirea constructului şi urmăreşte să verifice dacă eşantionul de stimuli şi cel de răspunsuri observate şi înregistrate în procesul de măsurare sunt reprezentative pentru universul de comportamente pe care îl defineşte constructul respectiv (M. M. Linehan, după Silva, 1993).

Nu se poate afirma că, în privinţa atributului măsurat de test, un subiect care a realizat un scor mai mare îi este superior unuia care a obţinut un scor mai mic. Deci, operaţiile cuprinse sub denumirea de „validarea testului relativă la conţinutul său" nu reprezintă o validare, în sensul definiţiei date de *Standardele* APA din 1985. Ele realizează doar analiza testului în privinţa clarităţii, a reprczentativităţii şi a relevanţei conţinutului său. Ar putea fi privite ca o componentă a validării testului.

Aurel Stan(2000, apud Mitrofan, N., 2000) menţionează că mai mulţi autori au propus diferiţi indicatori de măsurare a validităţii de conţinut. Astfel, Lawshe a propus următoarea formulă pentru calcularea unui coeficient de validitate de conţinut, CVR (iniţialele, în limba engleză, de la *content validity ratio*):

$$CVR= \frac{N_e - \frac{N}{2}}{\frac{N}{2}}$$

Formulă în care Ne= numărul evaluatorilor (experţilor) care consideră testul, respectiv itemul, ca fiind reprezentativ; N = numărul total de evaluatori (experţi)

Validitate de aspect(sau de faţadă)

Acest tip de validitate nu este de fapt o măsură reală a validităţii, în sens tehnic (în Standardele APA nici nu o includ printre celelalte tipuri de validitate). Se referă la aprecierea pe care o fac cei care constituie obiectul testării cu o anumita probă, legat de adecvarea acesteia pentru categoria respectivă de subiecţi.

L.Guttman şi S. Shye(Nevo, 1993, apud Albu, M., 1998) formulează o definiţie operaţională a validităţii de aspect: “Se face o măsurare a validităţii de aspect atunci când o persoană testată/ un utilizator neprofesionist/un individ interesat, utilizând o tehnică absolută/ relativă, apreciază un item/

un test/ o baterie de teste drept foarte potrivit/ potrivit, pînă la nepotrivit pentru utilizarea care îl interesează". Aprecierea validităţii de aspect nu este realizată de experţi şi pentru măsurarea sa nu se folosesc metode psihometrice. Un inginer căruia într-un test de vocabular i se includ cuvinte ca "facsimil", "unicorn", "sinapsa " etc., poate judeca testul respectiv ca nepotrivit profesiei sale. Testul, spunem, în acest caz, nu are validitate de aspect. Validitatea de aspect este importantă pentru că asigură o atitudine pozitivă atât faţă de test cât şi faţă de examenul psihologic. Ea este legată mai mult de motivaţia celui testat decât de aptitudinea măsurată, şi astfel este inclusă ca parte componentă în performanţa generală la examenul psihologic.
La estimarea validităţii de aspect contează în primul rând aprecierile subiecţilor şi în al doilea rând experţii participanţi la construcţia testului. Validitatea de aspect nu face apel la metode psihometrice /statistice de calcul, ea se bazeaza pe impresii şi opinii, culese imediat după testare.

Pentru psihologii antrenati în activităţi cu caracter aplicativ, construcţia de teste şi validarea lor, reprezinta acţiuni de mare importanţă . În acest domeniu experienţa acumulata este mare, deşi încă nu s-a spus ultimul cuvânt.

7.3. EROAREA STANDARD A ESTIMĂRII, SEE

Formula de calcul utilizată pentru eroarea standard a estimării este:

SEE = Sdy-rxy2

în formulă:

- Sdy - deviaţia standard la scorurile obţinute la criteriu;
- rxy2, - pătratul valorii coeficientului de validitate;
- dacă rxy = maxim 1, atunci SEE 0, deci o certitudine de 100% privind comportamentul subiectului în viitor;
- dacă rxy = 0.80, atunci SEE = Sdy- 0.64 =0 .36, deci criteriul folosit de cei care au utilizat testul este bun, adică are putere de anticipare.

De exemplu, dacă rxy = 0.80, SEE = 0.36, anticiparea poziţiei ocupate de subiect este mai precisă cu 36%) decât dacă s-ar face din pură întâmplare.
Formula SEE mai poate fi scrisă astfel:

$$SEE= \sigma \cdot \sqrt{1-r^2}$$

Unde σ este abaterea standard a valorilor criteriului, r este coeficientul de corelaţie intre valorile predictorului şi valorile criteriului.
Cu cât coeficientul de validitate al testului este mai mare cu atât eroarea standard a estimării este mai mică, iar predicţia realizată cu ajutorul testului este mai precisă.

În acest context se pune întrebarea :cât de mare poate fi valoarea coeficientului de validitate? Nu există un răspuns acceptat unanim de către autori. Desigur, ca o cerinţă generală, cu cât valoarea este mai mare cu atât mai valid este testul respectiv. Să nu uităm însă, că absolut întotdeauna valoarea coeficientului de validitate al unui test este mai mică decât valoarea coeficientului de fidelitate al aceluiaşi test.

Unii autori (Anastasi, A., 1988, 1996) susţin că valoarea coeficientului validităţii trebuie să fie destul de mare pentru a fi statistic semnificativă la un nivel acceptabil, cum ar fi 0.01 sau 0.05.

Alţi autori, cum este cazul lui M. Smith (apud Stan, A., 2002), propun o gradare a semnificaţiilor unui indice de validitate:

a) peste 0,50 - excelent;
b) între 0,40 şi 0,49 - bun;
c) între 0,30 şi 0,39 - acceptabil;
d) sub 0,30 - slab.

A. Stan atenţionează însă că un coeficient de validitate mai mare de 0,70 este excepţional şi că el merită o examinare atentă când este raportat.

Hull(Guilford, 1965, Albu, M, 1998) a impus condiţia ca un test să fie utilizat în practică doar dacă are un coeficient de validitate mai mare decât 0,45, o cerinţă greu de realizat. Rareori un test bun poate avea o corelaţie mai mare de 0,5 cu un criteriu important(Murphy, Davidshofer, 1991).

În experienţa obişnuită, valoarea coeficientului de validitate al testelor utilizate pentru a face predicţii este cuprinsă între 0 şi 0,60, în majoritatea cazurilor între 0 şi 0,30. Constatarea că, în general, comportamentul aceleiasi persoane variază de la o situaţie la alta a condus la concluzia că valoarea coeficientului de validitate al testelor de personalitate administrate în scopul predicţiei unei acţiuni nu poate fi mai mare decât 0,4(Funder, 1983, Albu, M., 1998)

7.4. ASPECTE PRIVIND INTERPRETAREA VALIDITĂŢII

Din practică a reieşit observaţia că un test ideal din punct de vedere al validităţii este cel ai cărui itemi măsoară fiecare un alt factor; deci un test cu consistenţă internă scăzută, în mod paradoxal, aceasta însemnând o fidelitate mică. Coeficientul de validitate al oricărui test nu poate depăşi coeficientul de fidelitate al acestuia. Astfel, dacă un test nu este fidel, validitatea sa relativ la orice criteriu nu poate fi mare. Cu toate acestea, o valoare mare a coeficientului de fidelitate nu garantează o validitate mare.
La fel ca şi în cazul fidelităţii, o serie de **factori afectează şi validitatea** unui test. Amintim câţiva dintre aceştia(Albu, M., 1998):

a. Unele caracteristici ale populaţiei testate:
    - Sexul
    - Vârsta
    - Pregătirea profesională
    - Anumite trasături de personalitate

Acestea acţionează ca variabile moderatoare, modificând relaţia dintre criteriu şi predictor. Subiecţii se împart în grupe, în funcţie de valorile variabilei moderatoare, iar valoarea coeficientului de validitate se poate schimba de la o grupă la alta. În general coeficienţii de validitate sunt mai mici în grupele mai omogene, adică în grupele în care scorurile testului sau valorile criteriului variază puţin. Atunci când valoarea coeficientului de validitate este scăzută, iar ceea ce se cunoaşte despre test şi despre criteriu nu oferă o explicaţie, trebuie să se verifice dacă nu există o variabilă moderatoare care se interpune între criteriu şi test. Deoarece validitatea unui test este dependentă şi de grupul de subiecţi, atunci când se constată că un test nou este valid faţă de un anumit criteriu, este necesar să se efectueze o validare suplimentară(cross-validation), folosind acelaşi criteriu, dar un alt eşantion de subiecţi, proveniţi din aceeaşi populaţie ca şi primul. Dacă şi în acest caz apar dovezi ale validităţii atunci se poate recomanda utilizarea sa în practică.

b. Validitatea relativă la criteriu este influenţată de lungimea testului. Prin adăugarea de itemi paraleli la test validitatea testului creşte.
c. Validitatea testului este dependentă şi de validitatea criteriului, ca măsură a variabilei la care se referă testul. Dacă valorile criteriului sunt influenţate de scorurile testului se spune că a avut loc o *contaminare a criteriului*, ceea ce duce la creşterea artificială a validităţii relative la criteriu a testului. Contaminarea criteriului se produce atunci când testul şi criteriul au itemi comuni sau în acele situaţii când valorile criteriului sunt rezultatul unor evaluări făcute de experţi, iar aceştia cunosc scorurile obţinute de subiecţi la test şi, intenţionat sau nu, sunt influenţaţi de ele în formularea aprecierilor.
d. Mărimea coeficientului de validitate este influenţată şi de frecvenţa relativă în populaţie a persoanelor care reprezintă caracteristica la care se referă testul.Atunci când comportamentul pe care îl măsoară sau îl prezice este foarte rar întâlnit(exemplu psihotismul sau suicidul), validitatea relativă la criteriu este de obicei scăzută.
e. Uneori criteriul utilizat duce la un coeficient de validitate scăzut:
    - Când criteriul are o fidelitate scăzută

- Intervalul de valori pentru criteriu este restrâns, neacoperind decât o parte din mulţimea valorilor posibile
- Distribuţia valorilor criteriului nu este normală
- Valorile criteriului nu sunt acordate sau notate corect, din cauza unor factori ce ţin de evaluatori(oboseală, neatenţie, incompetenţa sau părtinire), sau de dificultatea de măsurare a criteriului respectiv.

Aspecte importante la validarea unui test şi la interpretarea coeficienţilor de validitate

a. Validitatea unui test este dependentă de scopul pentru care acesta este folosit şi de populaţia căreia i se administrează testul. Un test nu este valid în orice situaţie de utilizare.
b. Există diverse metode pentru validarea testelor: analiza de itemi, calculul coeficienţilor de corelaţie predictor- criteriu, analiza factorială etc. Alegerea strategiei de validare trebuie să fie determinată de scopul pentru care se va utiliza testul şi de cerinţele formulate de cel care a solicitat examinarea psihologică.
c. Validitatea nu poate fi estimată dintr-un singur coeficient, ci se deduce din acumularea dovezilor empirice şi conceptuale. Acesta este motivul pentru care, începând cu Standardele americane din 1985, se recomandă utilizarea formulării : "dovezi ale validităţii relative la…" construct, criteriu, conţinut, în loc de "validitate relativă la …".
d. Strategiile de validare se aplică încă din faza de construcţie a testului, şi se repetă până când se obţine un test acceptabil în ceea ce priveşte lungimea, fidelitatea, validitatea şi uşurinţa de administrare, cotare şi interpretare. Procedeele de validare sunt următoarele:
   - Pentru testele ce măsoară constructe- se introduc în test două categorii de itemi, unii reprezentativi pentru construct, şi unii despre care se presupune ca nu au nici o legătură cu constructul. Sunt reţinuţi în final doar acei itemi care se aseamănă între ei în privinţa răspunsurilor obţinute de la subiecţi, se analizează conţinutul pentru a vedea dacă exemplifică teoria constructuluice trebuie măsurat şi dacă se deosebesc de itemii consideraţi nerelevanţi pentru construct.
   - Pentru testele ce vor fi validate doar relativ la conţinutul lor- se introduc doar itemi ce acoperă domeniul de conţinut al testului. După administrare se reţin doar itemii la care s-au obţinut răspunsuri asemănătoare. Se verifică dacă ei corelează cu un criteriu care se referă la acelaşi domeniu de conţinut, sau dacă răspunsurile obţinute la aceşti itemi se deosebesc semnificativ între două grupe constrastante de subiecţi.
   - Pentru testele utilizate în predicţii sau în decizii, pentru care se poate alege un criteriu de validare, după administrare se reţin itemii care se comportă faţă de criteriu aşa cum se doreşte să se comporte întregul test.
e. Validarea unui test se repetă mereu pe parcursul perioadei în care acesta este utilizat pentru îmbunătăţirea şi rafinarea interpretărilor scorurilor testului. La fel se procedează de fiecare dată când se fac modificări în orice aspect al testului sau al testării.
f. O serie de cercetători precum S. Messick(Foundations of validity; meaning and consequences in psychological assessment, apud Minulescu, M., 2003), susţin că ceea ce trebuie validat nu este numai testul sau instrumentul de observare ca atare, ci inferenţele derivate din scorurile la test. Aceste inferenţe privesc înţelesul scorurilor sau interpretării şi implicaţiile practice ce vor urma acestor interpretări. Astfel privită- validitatea se redefineşte ca: "o evaluare empirică a înţelesului şi consecinţelor măsurătorii efectuate. Validitatea combină cercetarea ştiinţifică cu

argumentul raţional(teoretic) pentru a justifica sau anula interpretarea unui test şi utilizarea lui.”

g. Fundamentul pentru consecinţele interpretării testului îl reprezintă evaluarea implicaţiilor înţelesurilor scorului, incluzând uneori implicarea denumirii constructului, teoriei lărgite care conceptualizează proprietăţile constructului şi ramificaţiile relaţionale ale înţelesului constructului şi ideologia mai largă care dă teoriei scop şi perspectivă(de exemplu ideologia referitoare la natura umană ca învăţare, adaptare etc.)
h. Deseori denumirea constructului relevă atitudinea latentă a cercetătorului faţă de importanţa acesteia, atitudine care este mai degrabă dependentă de ideologie decât de dovezile experimentale.
i. Validitatea de construct este acea forţă integrativă care uneşte problemele validităţii într-un concept unitar.

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition)*, Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition)*,Routledge, London and New York,
Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului,* Ed. Fundatiei Romania de Maine, Bucuresti,
Mitrofan Niculae- Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare*, Editura Polirom, Iaşi
Urbina Susana, *2004- Essential of psychological testing*, John Wiley & Sons, Inc. Hoboken, New Jersey

8. CONSTRUIREA ŞI EXPERIMENTAREA TESTELOR

8.1.Cerinţe: paşii standard în construirea unui test psihologic
8.2.Analiza de item

8.1. CERINŢE: PAŞII STANDARD ÎN CONSTRUIREA UNUI TEST PSIHOLOGIC
(Minulescu, M., 2003)

• **Stabilirea unei necesităţi:** luarea în considerare a cerinţelor formale şi informale pentru test. Trebuie să realizăm un fel de anchetă socială pentru a vedea dacă această necesitate este reală sau ţine de imaginarul nostru. Trebuie să scanăm disponibilitatea pieţei, să vedem dacă există ceva similar (pentru a nu repeta).

• **Definirea obiectivelor** şi **parametrilor testelor**: stabilim scopul testului (cine este testat şi de ce a fost ales acest test), clarificăm modul în care informaţia utilizată de test va fi utilă şi celui care dă testul şi beneficiarului. Trebuie să luăm în considerare tipul de format al itemilor şi tipul de format al răspunsurilor, dar şi numărul de itenii ce vor fi incluşi.

• **Selectarea unui grup de experţi** în domeniul respectiv: se discută din nou obiectivele, scopurile şi parametrii testului şi se determină o primă machetă a testului.

• **Scrierea itemilor(crearea băncii de itemi):** se utilizează experţi din domeniul testului sau specialişti din aria domeniului respectiv pentru a scrie itemii. După ce au fost sorişi, itemii trebuie să fie revizuiţi **(**din punctul de vedere al conţinutului) de către cel puţin o persoană sau de către o echipă care nu a fost implicată în scrierea lor.

• **Faza de teren.** Itemii sunt supuşi probei realităţii. Un prim eşantion de itemi este utilizat pentru a fi testaţi subiecţi reali, dintr-un anumit grup ţintă (grup realizat în funcţie de vârstă, sex, pregătire profesională etc.)- In urma testării, se va calcula gradul de dificultate şi de discriminare al itemilor (aceasta reprezintă analiza de itemi).

• **Revizuirea itemilor,** care se produce în urma analizei de itemi. Se verifică dacă există o anumită încărcătură nedorită de itemi sau prejudecăţi (de exemplu, sexuale sau legate de minorităţi). Sunt eliminaţi itemii care pot să fie incorecţi, nedrepţi sau pot să lezeze anumite grupuri de indivizi.

• **Alcătuirea formei finale a testului.** Se verifică adecvarea grilei de scorare, mergându-se până la subtilităţi în ceea ce priveşte scorarea răspunsurilor. Se întâmplă foarte multe accidente în faza finală, astfel că este necesară intervenţia unui „cap limpede" (cineva care este scos din paşii experimentării şi care verifică tot, inclusiv forma finală; de obicei, este cineva din afară). Se reverifică itemii şi grila de scorare.

• **Constituirea normelor şi calcularea finalităţii** şi **validităţii:**

1. Datele tehnice ale testului sunt verificate. Se fac procedurile de eşantionare sau de constituire a loturilor de experimentare.
2. Administrarea şi scorarea formei finale a testului.
3. Se calculează finalitatea si itemii de finalitate.
4. Construirea normelor adecvate de interpretare.

8.1.2. METODE DE CONSTRUIRE A TESTELOR

În majoritatea lucrărilor de teoria testelor psihologice, metodele de construire a testelor psihologice sunt împărţite în trei categorii(Burisch, 1986; Van Der Maesen, Hofstee, 1989; Albu M., 1998):

1. Metodele deductive- denumite şi raţionale, sau bazate pe simţul comun
2. Metodele inductive sau interne
3. Metodele externe sau empirice

1. **Metodele deductive**, denumite şi **raţionale sau bazate pe simţul comun**, sunt cele în care construirea itemilor şi selectarea celor care vor compune testul se bazează exclusiv pe părerile unor persoane (experţi sau nu). Mai întâi sunt alese constructele care vor fi măsurate de teste, apoi sunt elaboraţi itemii, care realizează o definire implicită sau explicită a constructelor,

singurul ghid în această operaţie fiind intuiţia autorilor itemilor.

Din această categorie fac parte:

- metoda prototipului;
- metoda BARS(behavioral anchored rating scales);
- metoda ordonării comportamentelor de către experţi;
- metoda designului cu faţete;
- metoda intervalelor aparent egale.

2. **Metodele inductive sau interne** pornesc de la o colecţie mare, eterogenă, de itemi (de obicei extraşi din teste aflate în circulaţie), considerată a fi reprezentativă pentru constructele care trebuiesc măsurate. Prin procedee statistice, cum sunt analiza factorială şi analiza de clusteri, se urmăreşte să se depisteze grupările de itemi care corespund constructelor respective.

Din această categorie de metode fac parte:

- metoda consistenţei interne;
- metoda teoretică;
- metoda analizei criteriului intern.

3. **Metodele externe** sau **empirice** necesită precizarea, de la început, a relaţiilor care vor trebui să existe între scorurile testului care va fi construit şi diverse criterii externe. Se porneşte de la o colecţie mare, eterogenă, de itemi, se compară scorurile acestora cu criteriile şi se reţin cei care au comportarea pe care trebuie să o aibă întregul test.

Din această categorie de metode fac parte:

- metoda criteriului extern;
- metoda analizei constructului.

După cum se observă, metodele deductive nu necesită nici experimentarea itemilor, nici efectuarea unor analize statistice, în timp ce metodele inductive şi cele externe se bazează pe prelucrarea statistică a rezultatelor experimentării itemilor, acordând o pondere redusă modului de construire a itemilor. În multe situaţii este posibilă - şi chiar recomandabilă - construirea testelor cu ajutorul mai multor metode: o metodă deductivă va servi la formarea băncii de itemi, iar printr-o metodă internă sau/şi una externă se vor selecta itemii care vor alcătui testul.

8.2. ANALIZA DE ITEM

8.2.1. ANALIZA GRADULUI DE DIFICULTATE ŞI A CAPACITĂŢII DE DISCRIMINARE A ITEMULUI

Analiza de itemi are ca obiectiv de bază descifrarea mecanismelor cognitive aplicate de subiecţi pentru formularea răspunsurilor la itemi şi verificarea calităţilor itemilor ca instrumente de măsură sau de predicţie. Pe baza analizei de itemi se obţin informaţiile care permit selectarea itemilor care intră în componenţa testului.

Itemii pot fi analizaţi dintr-o dublă perspectivă - cantitativă şi calitativă.

Analiza cantitativă se referă la proprietăţile statistice ale itemilor şi este focalizată în principal pe clarificarea problemelor privind ***dificultatea şi capacitatea de discriminare a itemilor.***

Analiza calitativă se referă în principal la aspecte de ***conţinut şi de formă*** incluzând problema evaluării eficienţei procedurilor de redactare şi a validării de conţinut.

Analiza itemilor poate fi rezumată prin următorul algoritm:

1) calculul indicelui de dificultate pentru toţi itemi şi eliminarea celor care sunt rezolvaţi de toţi subiecţii şi celor nerezolvaţi de nici un subiect;
2) depistarea cauzelor pentru care unii indici de dificultate sunt foarte mari sau foarte mici şi eliminarea itemilor cu greşeli;
3) în situaţia itemilor cu răspunsuri la alegere, se analizează răspunsurile incorecte şi se elimină aceia

în care unele răspunsuri greşite au fost alese de foarte mulţi sau foarte puţini subiecţi;

4) aplicarea uneia dintre procedurile analizei de itemi şi selecţia itemilor, în funcţie de condiţiile pe care trebuie să le îndeplinească testul ce se construieşte, respectiv:
   a. dacă testul trebuie să discrimineze între două grupe contrastante, respectiv să permită scoruri care diferă mult de la o grupă la opusul ei, se va calcula pentru fiecare item indicele de discriminare, eliminându-se itemii necorespunzători;
   b. dacă testul trebuie să se comporte într-o manieră prestabilită faţă de un anume criteriu, se vor elimina itemii necorespunzători;
      i. se vor elimina itemii cu indicele de dificultate necorespunzător scopului pe care urmează să-l îndeplinească testul. In principiu se vor elimina cei foarte uşori şi foarte dificili. Dacă cerinţa este ca lestul să identifice pe cei care prezintă nivele foarte scăzute şi foarte ridicate pentru o anume trăsătură sau facultate, se vor reţine doar aceste tipuri de itemi;
      ii. se va calcula coeficientul de corelaţie între scorurile testului şi scorurile itemului şi se vor elimina itemii care prezintă corelaţii nesemnificative sau negative.
5) In continuarea construirii testului pe baza itemilor astfel selectaţi se procedează la studierea caracteristicilor psihometrice (fidelitate, validitate, dificultate, putere de discriminare), aspectele legate de lungimea testului, tipuri de itemi din care este compus, timpul de rezolvare a testului (inclusiv variante privind standardizarea timpului), costurile administrării testului.

Dificultatea itemului

In majoritatea situaţiilor cerute de scopul testării, **dificultatea itemului e definită** în funcţie de procentul de persoane care răspunde corect la el. In procesul de construire a unui test, motivul principal al analizei dificultăţii itemilor este de a alege itemi care au un procent de dificultate adecvat, în aşa fel încât testul să nu constituie nici o dificultate de netrecut pentru majoritatea subiecţilor, nici să nu fie rezolvabil de marea lor majoritate.

- Pentru diferite loturi de subiecţi, cu caracteristici diferite privind, de exemplu, vârsta, nivelul de pregătire academică, tipul de pregătire, mediul de provenienţă etc, aceiaşi itemi pot conduce la diferite procente sau ponderi de subiecţi care răspund corect.
- Majoritatea testelor sunt construite cu itemi având grade de dificultate diferite astfel se pune problema asigurării unui nivel de dificultate optim şi a posibilităţii de a compensa interitemi gradul de dificultate.
- Nivelul de dificultate recomandabil este .50 (50% reuşită). Cu cât un item se apropie de 0% sau de 100%, cu atât este ineficient în diferenţierea subiecţilor. Itemul trebuie să fie capabil să diferenţieze între toţi cei care îl reuşesc şi cei care nu îl reuşesc pentru a avea valoare de informaţie diferenţială.

  De exemplu, pentru gradul de dificultate de 0.5 (50%): Să presupunem că din 100 de persoane, 50 reuşesc şi 50 nu reuşesc să rezolve itemul. Deci, itemul ne va da posibilitatea să diferenţiem între fiecare dintre cei care l-au reuşit şi fiecare dintre cei care au eşuat. Deci, avem astfel 50 x 50 = 2.500 perechi de comparaţie, sau biţi de informaţie diferenţială. Un item reuşit de .70 (70%) va favoriza 70 x 30 = 2.100 biţi informaţionali. Unul reuşit de .9 (90%) furnizează, deci, 90 x 10 = 900, iar cel reuşit de 100%, 100 x 0 = 0. Acelaşi lucru este valabil şi pentru itemi mai dificili, pe care îi reuşesc sub 50% dintre subiecţi.

  - Itemii din cadrul unui test tind să intercoreleze. Cu cât este mai omogen testul, cu atât gradul de intercorelare este mai mare. In situaţia extremă în care toţi itemii ar fi perfect intercorelaţi şi toţi ar avea gradul de dificultate .50, aceleaşi 50 de persoane din 100 vor rezolva fiecare dintre itemi. Deci, jumătate dintre subiecţi vor avea scoruri perfecte, iar cealaltă jumătate vor avea un rezultat nul. Deci, datorită intercorelării dintre itemii testului cel mai bine este să fie selectaţi itemi cu nivele de dificultate diferite a căror medie a dificultăţii să fie .50. Cu cât e mai mare nivelul de intercorelare dintre itemi, cu atât mai largă trebuie să fie gama de grade de dificultate a itemilor selectaţi.
  - În cazul scalelor de interval, procentul de persoane care reuşeşte un item exprimă dificultatea itemilor la nivelul scalei ordinale, ceea ce înseamnă că exprimă corect rangul şi dificultatea relativă a itemilor. De exemplu, dacă avem 3 itemi care sunt rezolvaţi, respectiv de 30%, 20% şi 10% dintre subiecţii lotului, putem conchide că primul item este cel mai uşor, iar itemul al treilea,

cel mai dificil, gradul de dificultate crescând de la primul la al treilea. Dar, pentru diferenţe de procentaje egale, nu putem aprecia dacă există şi diferenţe egale în gradul de dificultate între cei trei itemi. Acest lucru ar fi posibil doar în cazul unei distribuţii rectangulare, unde cazurile ar fi uniform distribuite pe tot şirul. Scorurile de tip percentil nu reprezintă unităţi egale, ele diferă în mărime de la centru la extremele distribuţiei.

- Dacă avem o distribuţie normală a rezultatelor pentru trăsătura respectivă, nivelul de dificultate al itemului poate fi exprimat în termenii unei scale unităţi de interval egale comparativ cu tabelul de frecvenţe al curbei normale. In acest caz, avem informaţia că 34% din populaţie intră în zona cuprinsă între medie şi - sau + 1 sigma(σ).

  Astfel, un item care are nivelul de dificultate .84 (rezolvat de 80% dintre subiecţi) va include jumătatea superioară (50%) plus 30% din cazuri din jumătatea inferioară (50 +30 = 80). Deci, itemul cade 1 sigma sub medie. Un item reuşit de 16% dintre subiecţi va cădea la o sigma peste medie (peste acest punct se află 16% dintre cazuri, respectiv 50 - 34 (o sigma) =16. Un item reuşit de exact 50% dintre subiecţi cade pe medie şi va avea valoarea 0 pe această scală.

  In practică, datorită faptului că dificultăţile itemilor exprimate în termenii distanţelor pentru o curbă normală include valori pozitive şi negative, Serviciul american de testare educaţională a propus următoarea ecuaţie de convertire:

Δ (delta) = 13 + 4z

Constantele 13 şi 4 au fost alese arbitrar pentru a permite o scală care să elimine valorile negative. Z indică numărul de deviaţii standard(abateri standard) de la medie (sigma).

De exemplu, un item reuşit de aproximativ 100% subiecţi va cădea la o distanţă de 3 sigme de medie, deci va avea D = 13 + 4x (-3) = 1.

La cealaltă extremă, un item cu un procent de reuşită sub 1 va cădea la + 3 sigme şi va avea D = 13 + 4x (3) = 25. Un item care cade pe medie are D = 13 + 4x (0) =13. Deci toţi item ii vor cădea la scale D între 1 şi 25, iar dificultatea medie va fi redată prin nivelul 13.

Analiza capacităţii de discriminare a itemilor

Discriminarea itemilor se referă la gradul în care un item diferenţiază corect întresubiecţi în ceea ce priveşte comportamentul destinat să-1 măsoare. în literatura de specialitate apar peste 50 de indicatori de discriminare care pot fi utilizaţi în construirea diferitelor tipuri de teste, indicatori care, de regulă, oferă rezultate relativ asemănătoare.

O metodă utilă pentru grupe în general mici cuprinde următoarea procedură:

- Sunt pretestaţi subiecţii unui lot restrâns (ex. 60 de persoane).
- în funcţie de rezultate, subiecţii sunt grupaţi în trei clase. 20 cu scorurile cele mai înalte (clasa de sus, „U" de la *upper)*, 20 cu scorurile cele mai scăzute (clasa de jos, „L" de la *lower)*, şi 20 cu scoruri intermediare (clasa medie, „M", de la *middle).*

Pe cele 3 clase se vor verifica itemii testului astfel:

Tabelul ni: I

Item	U	M	L	Dificultatea	Discriminarea
				U+M+L	U-L
1	15	9	7	31	8
2	20	20	16	**56***	4
3	19	18	9	46	10
4	10	11	16	37	**-6***
5	11	13	11	35	**0***
6	16	14	9	39	7
7	5	0	0	**5***	5
etc.					

Dificultatea itemului reprezintă suma de reuşite la cele trei clase de subiecţi, iar **discriminarea** este **dată de diferenţa** dintre grupele extreme.

În aceste condiţii, putem vedea că **există 4 itemi care prezintă probleme** (*) fie din perspectiva dificultăţii, itemii 2 şi 7, **fie** a discriminării, itemii 4 şi 5. Dacă itemul 2 are o dificultate prea mică, 7 este prea dificil, deci trebuie excluşi. Itemii 4 şi 5 au discriminare fie negativă, fie nulă, deci vor fi excluşi. De regulă, în situaţia claselor de subiecţi care au aceeaşi dimensiune, itemii cu valoare de discriminare scăzută sunt cei de la 3 puncte în jos.

Indexul de discriminare

Când exprimăm în procente numărul de subiecţi care reuşesc la itemii cuprinşi într-un nou test, o diferenţă de 2% reprezintă indexul de discriminare, indiferent de dimensiunea grupelor. Acest index de discriminare este denumit ***Upper-lower discrimination,*** prescurtat ca „U-L", „ULI", sau „ULD" sau pur si simplu „D".

De exemplu, calcul D pornind de la datele anterioare se prezintă astfel:

Tabelul nr. 2

Item	Procentaj de reuşită		Index de discriminare
	Clasa U	Clasa L	(diferenta)
1	75	35	40
2	100	80	20
3	95	45	50
4	50	80	-30
5	55	55	0
6	80	45	35
7	25	0	25
etc.			

D poate avea o valoare cuprinsă între +100 şi - 100.

Dacă toţi subiecţii clasei U reuşesc la item şi nici unul dintre subiecţii clasei L nu reuşeşte, D = 100. Dacă nici un subiect din U nu reuşeşte şi toţi cei din L reuşesc, avem valoarea lui D = 0.

Pentru unii dintre indicii de discriminare, valorile lui D nu sunt independente de dificultatea itemului, dar sunt influenţate în direcţia nivelului de dificultate intermediară.

Tabelul nr.3 indică maximul posibil al valorii lui „D" în funcţie de diferite procente de răspunsuri corecte. Dacă fie 100% fie 0% din lot reuşesc la un item, nu apare nici o diferenţă între clasele de subiecţi, deci „D" este 0. Dacă 50% reuşesc un item, este posibil ca toţi cei din clasa „U să-1 treacă, şi nimeni din clasa „L, iar „D" va fi 100 - 0 == 100. Dacă 70% reuşesc, maximul valorii pe care o poate lua „D" va fi 60 pentru că „U" 50/50= 100% şi „L" 20/50 = 40%. „D" va fi 100 - 40 = 60.

Pentru majoritatea scopurilor de testare, sunt preferabili itemii cu dificultatea 50%. Indicii de discriminare care favorizează acest nivel de dificultate vor fi adecvaţi pentru selecţie.

Relaţia dintre valoarea maximă a lui „D" şi dificultatea itemilor:

Tabelul nr. 3

Procentul de reuşită la item	Valoarea maximă a lui D
100	0
90	20
70	60
50	100
30	60
10	20
0	0

8.2.2. TEORIA RĂSPUNSULUI LA ITEM (T.R.I.)

În literatura de specialitate teoria răspunsului la item nu este considerată o teorie în adevăratul sens al cuvântului, ci o colecţie de modele şi metode statistice cu ajutorul cărora se dă un sens datelor provenite din măsurători psihologice(Albu, M., 1998).

Teoria răspunsului la item a fost denumită şi teoria trăsăturii latente şi teoria curbei caracteristice (TCC). Aspectul fundamental al acestei abordări este faptul că performanţa la item este legată de cantitatea estimată de „răsătură latentă" a celui care răspunde, reprezentată prin Θ*(theta).* Denumirea de trăsătură latentă se referă la un construct statistic (ceea ce nu implică automat că ar exista o entitate psihologică sau fiziologică corespunzătoare, cu o existenţă independentă). De exemplu, în testele

cognitive, trăsătura latentă este reprezentată, de obicei de abilitatea măsurată de test, iar scorul total la test este considerat adesea ca reprezentând o estimare iniţială a acestei abilităţi.
Diferite modele T.R.I. utilizează diferite funcţii matematice, bazate pe seturi de presupuneri diferite, dar rezultatele obţinute prin aceste modele sunt substanţial asemănătoare.

T.R.I. are la bază următoarele trei postulate:

I. Comportamentul unui subiect la un item al unui test poate fi prezis cu ajutorul unui set de factori, denumiţi trăsături latente (orice construct inobservabil, presupus continuu, despre care o teorie psihologică afirmă că deosebeşte persoanele între ele poate ti privit ca factor, iar fiecărui factor i se asociază o variabilă cu valori numerice între - infinit şi + infinit, denumită variabilă latentă).
II. Putem afirma existenţa unui factor doar după ce se observă că răspunsurile la itemii prin care se doreşte măsurarea constructului, covariază.
III. Relaţia dintre performanţa la item a subiecţilor şi fiecare dintre trăsăturile care au legătură cu performanţa poate fi descrisă prin câte o funcţie crescătoare, denumită funcţie caracteristică a itemului sau curbă caracteristică a itemului. Funcţia arată cum depinde probabilitatea de a răspunde corect la item de nivelul trăsăturii.

O trăsătură nu poate fi nici observată, nici măsurată direct. Evaluarea ei cantitativă este indirectă, prin intermediul unor variabile observabile (de exemplu, răspunsul la itemii unui test) despre care se presupune că reprezintă adecvat variabila latentă.

Operarea cu variabile latente presupune următoarea procedură în trei etape:

1. Definirea variabilei latente în termeni operaţionali sau construirea variabilei. Acest pas cere identificarea acelor diferenţe individuale care fac posibilă reprezentarea fiecărui individ printr-un punct pe axa reală (variabilă unidimensională). Dacă reprezentarea cere utilizarea unui spaţiu cu mai multe dimensiuni, variabila este multidimensională.
2. Definirea situaţiilor controlabile (crearea cadrului de observaţii). Situaţiile sunt reprezentate de itemii testului, iar rezultatul observaţiilor este exprimat numeric prin scorul la test.
3. Alegerea modelului matematic pentru obţinerea variabilei latente pe baza valorilor variabilelor. Modelul matematic ales trebuie să transforme o mulţime discretă în mulţimea numerelor reale. Variabilele latente apar în două tipuri de modele prin care sunt reprezentate legăturile dintre variabilele psihologice: modele care măsoară cauzele unde se presupune că variabila latentă este cauza variaţiilor variabilelor observate; modele care măsoară efectele care presupun că variabila latentă este cauzată de variabilele observate.

Dacă se utilizează mai multe instrumente de măsură (variabile observate) pentru investigarea aceleiaşi variabile latente, se presupune că în spatele covariaţiilor existente între variabilele observate stau relaţiile dintre ele şi variabila latentă pe care o măsoară. Acest fapt stă la baza utilizării analizei factoriale ca metodă care permite determinarea valorilor unuia sau mai multor factori (variabile latente) pornind de la mai multe variabile observate (teste).
In situaţia itemilor dihotomici, putem ilustra consecinţa T.R.L presupunând, de exemplu, că toţi itemii testului măsoară aceeaşi variabilălatentă unidimensională. T.R.I. impune, în această situaţie, ipoteza unidimensionalităţii spaţiului variabilelor latente, respectiv faptul că nu există o altă variabilă latentă, pe lângă cea specificată, care să influenţeze scorurile itemilor. Această ipoteză asigură îndeplinirea condiţiei de independenţă locală a itemilor, care cere ca modul în care o persoană răspunde la un item să nu fie influenţat de felul cum a răspuns la ceilalţi itemi. Conform T.R.I., se va construi, pentru fiecare item al testului,funcţia caracteristică a itemului care leagă valorile variabilei latente măsurate de item cu scorurile obţinute la test, respectiv $P(y)$: R - > [0,1]. Probabilitatea ca o persoană având nivelul variabilei latente măsurate de item egală cu y să răspundă corect la item, pentru orice valoare a lui y aparţinând realului. Cu ajutorul acestei funcţii se pot aprecia calităţile itemului, respectiv dificultatea şi puterea de discriminare, iar graficul acestei funcţii vizualizează comportarea itemului ca instrument de măsură. T.R.I. introduce ipoteza că funcţia P este crescătoare. Dacă itemul este un instrument bun de măsură, un subiect care are un nivel mai înalt al variabilei latente va avea şanse mai mari de a rezolva corect itemul decât cel care are un nivel mai scăzut.

Scopul teoriei răspunsului la item este de a elabora metode de estimare a valorii variabilei latente la subiecţii testaţi şi metode de estimare a caracteristicilor itemilor testului, pe baza răspunsurilor dale de subiecţi la itemi.

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice*, Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition)*, Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition)*,Routledge, London and New York,
Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului*, Ed. Fundatiei Romania de Maine, Bucuresti,
Mitrofan Niculae- Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare*, Editura Polirom, Iaşi
Urbina Susana, *2004- Essential of psychological testing*, John Wiley & Sons, Inc. Hoboken, New Jersey

9. ETALONAREA ŞI REETALONAREA TESTULUI

9.4. Norme: definiţii
9.5. Eşantionul normativ, cerinţe privind eşantionarea
9.6. Tehnici de normare

9. 1. NORME: DEFINIŢII

Obiectivul major al metodei statistice în psihodiagnoză este acela de a organiza şi sistematiza datele cantitative pentru a permite o interpretare obiectivă şi clară. Rezultatul, **nota brută** obţinută la test pentru a putea fi interpretată, trebuie să aibă sens şi consistenţă.

Normele reprezintă performanţele realizate la testare de către un eşantion luat ca model; astfel, normele sunt stabilite empiric - prin determinarea a ceea ce se poate obţine prin testare din partea unui grup reprezentativ (raportăm fiecare rezultat individual la rezultatele obţinute de eşantionul standard ca să putem spune cum se încadrează individul în cadrul rezultatelor populaţiei - dacă acestea coincid cu rezultatele medii ale grupului, sunt sub acestea sau peste acestea).

Pentru determinarea exactă a locului persoanei în această distribuţie de scoruri, rezultatul este transformat într-o unitate de măsură normată. Aceste valori derivate au un scop dublu:

- precizează locul individului faţă de alţi indivizi din eşantion;
- permit compararea între rezultatele individuale la diferite teste. Există modalităţi diferite de a transforma rezultatele brute la un test pentru a corespunde celor două scopuri.
- în general, rezultatele derivate sunt de trei tipuri:
    - norme de vârstă,
    - percentile şi
    - scoruri standard

9.2. EŞANTIONUL NORMATIV. CERINŢE PRIVIND EŞANTIONAREA

Orice normă se limitează la o populaţie normativă pe baza căreia a fost derivată; cel care foloseşte testul şi modul de normare trebuie să cunoască modul cum s-au stabilit aceste norme. In testarea psihologică, normele nu au valoare permanentă şi absolută, ci reprezintă performanţa la testare a subiecţilor care constituie eşantionul de normare. In alegerea eşantionului, efortul este de a stabili un grup reprezentativ şi variat pentru populaţia căreia îi este destinată proba. In alegere se fac eforturi pentru reprezentativitate. Există o serie de factori care fac ca eşantionul să nu fie reprezentativ. Legată de problema reprezentaţivităţii este aceea de a defini populaţia specifică la care se aplică normele. În mod ideal, ar fi necesar să se definească populaţia pe baza obiectivelor testului, dar sunt dificultăţi practice în obţinerea subiecţilor care fac irealizabil acest deziderat.
In practică, marea majoritate a testelor nu sunt standardizate pe populaţii extinse. Eşantionul pentru normare trebuie să fie suficient de mare pentru a asigura valori stabile, astfel încât un alt eşantion extras din aceeaşi populaţie să nu producă norme care să difere semnificativ. Se cere ca un alt eşantion din aceeaşi populaţiei să nu producă rezultate diferite de rezultatele eşantionului iniţial.
Normele specifice, cel mai des întâlnite în practică, constituie standardizarea testelor pe o populaţie atent definită, aleasă pentru a se potrivi scopurilor specifice ale fiecărui test. In astfel de cazuri se vor raporta clar la norme limitele populaţie informative, specificându-se restricţia în aplicare. Pentru diversele scopuri ale testării sunt necesare norme foarte specifice.
în practică, chiar când există **norme reprezentative** pentru o populaţie definită în linii generale, sunt construite şi utilizate **şi norme de subgrup.** Acest lucru se întâmplă mai ales când subgrupele produc diferenţe suficient de mari la test (exemplu, populaţia uzinală dintr-un sector faţă de alt sector, o arie geografică faţă de alta). Subgrupele se pot constitui în funcţie de criterii precum: vârsta, fundalul cultural şi social, tipul de programe şcolare, sexul, regiunea geografică, mediul rural sau urban de provenienţă etc. Modul de utilizare a testului determină în ultimă
instanţă tipul de diferenţieri - criterii în funcţie de care utilizăm norme generale sau particulare.

Un caz particular este cel al **normelor locale,** produse când testul este folosit pentru un anumit mediu specific. Grupurile de normare folosite vor fi deosebit de minuţios pregătite şi definite.

Grupul fix de referinţă. Un tip de scală normativă utilizează un grup de referinţă fix pentru a putea asigura compatibilitatea şi continuitatea rezultatelor tară a asigura şi evaluarea normativă a performanţei. In astfel de situaţii, rezultatul individual va depinde de caracteristicile grupului testat într-o anume situaţie. Scalele constituite pe baza unui grup fix de referinţă sunt analoge celor de măsurători fizice.

Există şi situaţii mai speciale, când domeniul de referinţă nu este o populaţie de persoane, ci un domeniu al conţinutului. Acest tip de scale normative dau sens conţinutului rezultatelor, iar norma este definită ca un număr ce indică procentul unui eşantion organizai dintr-un domeniu definit de probleme pe care individul le-a rezolvat corect, de exemplu, mărimea vocabularului unei persoane.

Facem o diferenţă între **eşantion** care se referă la grup efectiv de oameni testaţi, şi **populaţie** - care reprezintă tipul de indivizi din care este selectat eşantionul pe baza criteriilor de reprezentativitate. De exemplu, în încercarea de a defini criteriile pentru elaborarea unui eşantion pentru o populaţie reprezentativă pentru elevul de liceu, ne putem întreba:
Cum să definim populaţia specifică la care se aplică normele: pe baza obiectivelor testului? Ce tip? Se aplică unor absolvenţi de liceu sau unui anumit nivel de dezvoltare a personalităţii în adolescenţă? Normele să fie semnificative pentru nivelul de vârstă şi cu aplicabilitate internă? Lărgită la nivelul unui oraş? La nivelul unei regiuni sau al întregii populaţii? Ideal este pentru un test să fie construit şi normat pe baza unui eşantion din populaţia generală a unei ţări: cu considerente legate de sex, vârstă, formare culturală, clase, diferenţe sat - oraş etc. (vezi date de recensământ).

Etalonarea este una dintre etapele acţiunii de standardizare a testului. Înseamnă construirea etaloanelor(a normelor).
Ea se desfăşoară în următorii paşi:
(1)Se defineşte populaţia pentru care se etalonează testul.

(2)Se extrage un eşantion din această populaţie.
(3)Se administrează testul acestui eşantion.
(4)Se determină formula de obţinere a cotelor transformate dinscorurile testului.

În continuare vom detalia conţinutul celor patru paşi.

(1) Când se construieşte un test, este important să se decidă cine sunt persoanele care vor fi examinate cu el, ce caracteristici are cel mai mare grup din care vor proveni ele. Acesta reprezintă *populaţia* căreia îi este destinat testul. Definirea populaţiei se poate face fie concret, prin enumerarea tuturor membrilor ei, fie operaţional, prin indicarea unui set de reguli care vor permite să se cunoască, pentru orice persoană, dacă aparţine sau nu populaţiei.

De exemplu, populaţia poate fi formată din "toate persoanele care au vârsta cuprinsă între 14 ani împliniţi şi 18 ani neîmpliniţi, care locuiesc în oraşul Bucureşti şi sunt elevi la liceu".

Definiţia populaţiei pentru care se face etalonarea unui test este foarte arbitrară: toţi indivizii sau numai un lot cu acelaşi nivel şcolar sau de aceeaşi vârstă, dintr-un oraş, o regiune sau o ţară.

(2) Din populaţie se va extrage un eşantion de indivizi -numit ***grup de normare***-căruia i se va administra testul, pentru a stabili normele.

Este important ca în eşantionul format testul să se comporte la fel ca în întreaga populaţie: să măsoare sau să prezică aceeaşi variabilă, iar repartiţia scorurilor la test în eşantion să fie asemănătoare celei care s-ar obţine dacă s-ar dispune de scorurile întregii populaţii.

Pentru aceasta, eşantionul constituit trebuie să îndeplinească două cerinţe:

•să fie reprezentativ pentru populaţia din care a fost extras (structura sa să fie asemănătoare structurii populaţiei, în privinţa zonelor geografice şi a mediului de domiciliu, a nivelului socio-economic, a vârstei şi a altor caracteristici pe care autorul testului le consideră importante şi care ar putea influenţa rezultatele testului);

•să fie destul de mare, aşa încât rezultatele obţinute la test să reflecte repartiţia acestor valori în populaţia din care provine.

Obţinerea unui eşantion

Se poate face prin **selecţie aleatoare simplă sau prin selecţie aleatoare stratificată.**

a. In primul caz, **fiecare membru al populaţiei are şanse egale** de a intra în componenţa eşantionului. Dacă eşantionul este numeros, atunci în el se vor găsi persoane aparţinând tuturor zonelor geografice, claselor sociale, naţionalităţilor prezente în populaţie, cu aceeaşi frecvenţă relativă ca în populaţie, deci eşantionul va fi reprezentativ pentru populaţie.

b. Atunci când volumul eşantionului este mic, pentru obţinerea unui eşantion reprezentativ este necesar să se facă o **selecţie aleatoare stratificată**.

În acest scop:

1. Se determină **variabilele de mediu** (sex, vârstă, rasă, clasă socială, nivel educaţional etc.) care au o influenţă mare asupra rezultatelor aplicării testului.
2. Se face apoi o **clasificare a populaţiei** după valorile acestor variabile şi
3. Se stabileşte care trebuie să fie compoziţia eşantionului astfel încât **fiecare clasă să fie reprezentată în eşantion proporţional** cu frecvenţa ei relativă în populaţie.
4. Din fiecare clasă **se extrag, prin selecţie aleatoare simplă**, persoanele care vor intra în componenţa eşantionului.

Principala dificultate în cazul selecţiei aleatoare stratificate constă în deteminarea variabilelor după care se face împărţirea în clase. Se recomandă (Kline, 1993) ca ele să se aleagă dintre variabilele care corelează semnificativ cu scorurile testului. Numărul lor nu va fi prea mare, pentru a nu se forma prea multe clase, ceea ce ar necesita alcătuirea unui eşantion foarte numeros. In general nu se aleg mai mult de patru variabile. Frecvent, variabilele după care se face clasificarea sunt sexul, vârsta şi clasa socială.

Volumul eşantionului se stabileşte astfel încât **din fiecare clasă să se extragă cel puţin 300** de persoane (Kline, 1993, 2000). Se deduce de aici că populaţiile generale, care sunt eterogene şi din acest motiv necesită utilizarea mai multor variabile pentru formarea claselor, impun obţinerea unor eşantioane foarte numeroase. Pentru a se face economie de timp şi de bani, adesea se renunţă la

stabilirea normelor pentru o populaţie generală şi se utilizează populaţii mai puţin numeroase şi mai omogene, pentru care eşantioanele pot fi de volum mai redus. în acest fel se obţin *norme locale ale testului.*
Dacă se constată că un subgrup al eşantionului, care poate fi identificat printr-o valoare a unei variabile de mediu (de exemplu "femei" sau "persoane din mediul urban") realizează la test scoruri mult mai mari sau mult mai mici decât restul eşantionului, se vor determina *norme ale subgrupului* şi acestea vor fi folosite în practică atunci când persoanele al căror scor trebuie interpretat au aceeaşi valoare a variabilei de mediu ca şi subgrupul.
(3)Dacă eşantionul are un volum foarte mare ori este format din persoane care domiciliază sau lucrează în locuri diferite, testul nu poate fi administrat deodată tuturor indivizilor. Examinarea cu ajutorul testului se va face pe subgrupe. Dar, pentru ca rezultatele să nu fie distorsionate, condiţiile de administrare trebuie să fie identice pentru toate subgrupele.
(4)În urma analizei scorurilor realizate la test de persoanele din eşantion se stabilesc normele testului. Apoi se determină modul în care se obţin cotele transformate din cotele brute. În acest scop se folosesc mai frecvent două tipuri de transformări:
• *transformări liniare,* care schimbă doar media şi abaterea standard **a** rezultatelor, păstrând nemodificată repartiţia cotelor, deci ordinea indivizilor din eşantion. Prin astfel de transformări se obţin **cotele standard** z şi **cotele standardizate;**
• *transformări neliniare,* sau *transformări de arie,* care schimbă repartiţia cotelor brute, facând-o să semene cu repartiţia uniformă (în cazul **centilelor, decilelor** sau **al cuartilelor)** ori cu cea normală (în cazul când se urmăreşte obţinerea unei **scale normalizate** cu 5, 7 sau 9 trepte). Acest fel de transformări schimbă unitatea scalei, astfel încât frecvenţa relativă a cotelor brute care ocupă un interval de valori particular va fi egală cu aria suprafeţei cuprinsă între intervalul transformat, desenat pe axa absciselor, şi:

• curba repartiţiei uniforme, pentru rangurile centile, decile şi cuartile
sau
• curba repartiţiei normale, pentru clasele scării normalizate.

Cotele transformate rezultate prin transformări liniare sau neliniare indică poziţia relativă a subiectului într-o populaţie.
Un alt tip de transformări utilizate în etalonare sunt cele care conduc la **clase echivalente** sau **la vârste echivalente.** Ele se aplică pentru teste care măsoară variabile fiziologice sau psihologice ale căror valori cresc cu vârsta (de exemplu înălţimea, greutatea, bogăţia vocabularului, performanţa la citire etc). Aceste cote transformate descriu performanţa subiectului sub forma indicării unei populaţii a cărei mediană sau medie a cotelor brute este aproximativ egală cu cota brută a persoanei respective.
Unele teste de inteligenţă utilizează pentru interpretarea scorurilor, în locul vârstelor echivalente, normele de vârstă mintală.

Observaţie: în România nu există norme generale bazate pe eşantioane pe întreaga populaţie a ţării; în SUA, testele Minnesota, Weschler, California, testele educaţionale au fost construite pe baza eşantioanelor reprezentative pentru întreaga populaţie a ţării. La noi Testul California s-a lucrat şi s-au obţinut normări doar pe loturi nereprezentative din armată, industrie, populaţia de studenţi; la Cluj, H. Pitariu a construit un etalon pentru populaţia Transilvaniei. Nu există date de eşantionare a populaţiei generale a ţării. Norme generale de interpretare s-au realizat doar pentru testul *screening* Denver.
In restul situaţiilor s-a lucrat cu norme specifice, obţinute pe eşantioane foarte limitate la grupuri de vârste, nivel intelectual, profesii, zone, unităţi industriale sau şcolare etc.
O altă problemă este cea a testelor de personalitate care au norme străine şi astfel dau diferenţe în interpretare, diferenţe de structură mentală şi culturală de care trebuie să ţinem seama.

9.3. TEHNICI DE NORMARE

1. Normele pe nivele de vârstă

Conceptul de vârstă mentală a fost introdus în revizia scalelor Binet-Simon în 1908. Probele individuale sunt grupate pe nivele de vârstă. De exemplu, probele pe care majoritatea copiilor de 8 ani le-au putut rezolva au fost cuprinse în testul dedicat pentru 8 ani. Rezultatul unui copil la acest test va corespunde celui mai înalt nivel de vârstă pe care este în stare să îl rezolve. Dacă un copil de 8 ani reuşeşte la probele cuprinse pentru nivelul de 10 ani, vârsta sa mentală este de 12, deşi cea cronologică este de 10. Este, deci, cu 2 ani înaintea vârstei sale, realizând performanţele unui copil de 10 ani. Rezultatele medii obţinute de copii în cadrul grupelor de vârstă reprezintă normele de vârstă pentru un astfel de test.
Dar acest gen de normare - ca normă în sine - corespunde foarte mult empiricului şi mai puţin statisticii. Pentru că, în practică rezultatele realizate de individ la teste, scale de vârstă de acest tip prezintă un grad de împraştiere destul de mare. Dacă un subiect are reuşite superioare vârstei sale pentru o serie, poate nu reuşeşte pentru altele, care sunt sub nivelul său de vârstă. Se introduce conceptul de vârstă de bază - nivelul cel mai înalt până la care se pot rezolva testele, mai jos de care toate testele pot fi corect rezolvate.

Un copil poate avea nivelul de bază pentru 11 ani, dar în plus există dimensiuni de performanţă exprimate în luni, care se adaugă la dimensiunea de bază - pentru acele teste care prezintă nivel de vârstă imediat supraordonat (dacă subiectul a reuşit la toate pentru 11 ani - are
11 ani, dar a mai reuşit şi din cele pentru 12 ani la câteva are 11 ani şi 2, 4, 9 luni, funcţie de câte peste 11 ani a reuşit).

Copilăria şi pubertatea sunt perioade în care inteligenţa poate fi scalată pentru că este în dezvoltare. Unitatea de vârstă mentală nu rămâne constantă cu vârsta, ci, o dată cu înaintarea în vârstă, tinde să se micşoreze. Un copil care la 4 ani este în retard cu aproximativ un an, când va împlini vârsta de 12 ani el se va afla în retard cu 3 ani. Un an de dezvoltare mentală la 3-4 ani este echivalent cu 3 ani de dezvoltare pentru vârsta între 9-12 ani.
Dezvoltarea mentală progresează mai rapid la vârste mici şi mai lent la vârste mai mari; pe măsură ce persoana se apropie de anii de maturitate unitatea de vârstă mentală se va micşora şi ea o dată cu vârsta. Unitatea de vârstă mentală nu rămâne constantă cu vârsta, ci tinde să se micşoreze pe măsură ce se înaintează în vârstă. Pentru a permite o interpretare uniformă, indiferent de subiect sau de vârstă, a fost introdus coeficientul de inteligenţă, Q.I. Stern şi Kuhlman sunt primii care subliniază necesitatea introducerii acestui sistem de măsurare, care a fost în practică utilizat prima dată pentru scalele Stanford-Binet.
Q.I.-ui reprezintă raportul dintre vârsta mentală şi cea cronologică, fracţia fiind multiplicată cu 100. Dacă vârsta mentală este aceeaşi cu vârsta cronologică, coeficientul de inteligenţă = 100. Rezultatul, reprezentat printr-un coeficient sub 100, indică gradul de distanţă faţă de normal; ca şi cel peste 100, reprezentând avansul faţă de normal. Pentru a putea compara direct valorile Q.I. la vârste diferite, deviaţia standard nu trebuie să varieze în funcţie de vârstă. Condiţia presupune ca valorile deviaţiei standard a vârstei mentale să crească proporţional cu vârsta. Folosirea normelor de tip Q.I. trebuie precedată de o verificare atentă a variabilităţii întâlnită la diferite vârste pentru a asigura condiţia de variabilitate uniformă a deviaţiei standard sau a creşterii proporţionale a vârstei mentale.
O altă limitare a rezultatelor normate în funcţie de vârstă este legată de faptul că nu pot fi utilizate decât pentru acele funcţii psihice care prezintă o schimbare clară şi consecventă o dată cu vârsta. Trăsăturile psihice care au o slabă legătură cu vârsta nu pot fi măsurate cu astfel de norme.

2. Normele exprimate prin percentile

Percentilul se referă la procentul de persoane dintr-un eşantion standard care se situează sub un rezultat dat. Dacă 30% dintre subiecţi rezolvă mai puţin de x itemi corect, acest rezultat neprelucrat, x, va corespunde celui de al 30-lea percentil. Gama de percentile este între 1 şi 99 şi reprezintă procentul de persoane din grupul de normare care au scoruri la nivelul sau sub nivelul unui anumit scor. Un percentil indică populaţia relativă pentru individ dintr-un eşantion standard. Percentilele pot fi considerate **ca ranguri într-un grup de 100 de subiecţi**.
Modul de aflare a percentilelor:

Se socotesc toate scorurile cu valori mai mici decât scorul care ne interesează; se împarte la numărul total de scoruri; se înmulţeşte cu 100. În rangul de percentile, scorul mediei este percentila 50, P 50, ca măsură a tendinţei centrale. Percentilele peste 50 reprezintă, succesiv, realizări deasupra mediei, iar cele sub 50, realizări inferioare. Percentilul 25 şi percentilul 75 sunt denumite primul şi al treilea **quartil** (sfert), delimitând sferturile inferior şi superior ale distribuţiei scorurilor în populaţia de referinţă. Rezultatul neprelucrat inferior oricărui rezultat din eşantionul standard şi rezultatul superior oricărui rezultat din eşantionul standard vor avea rangul P0 şi, respectiv, P 100. Exprimarea semnificaţiei scorului este de genul: Scorul persoanei este mai mare decat ex. 50%, 55%, 38% etc. dintre scorurile obţinute de celelalte persoane.

Avantajele rezultatelor exprimate în percentile:

- sunt uşor de evaluat şi au sens şi pentru nespecialişti;
- pot fi folosite şi pentru adulţi şi pentru copii, şi pentru orice tip de iest.

Dezavantajul principal constă în inegalitatea evidentă dintre unităţi, mai ales în zonele extreme ale distribuţiei (diferenţele dintre rezultatele brute sunt exagerate prin transformarea în procente; diferenţele dintre rezultatele aflate în apropierea extremităţilor sunt mult mai reduse).

3. Norme/scoruri standard

Rezultatele standard se pot obţine prin transformări lineare sau prin transformări nelineare ale rezultatelor brute.

Când se obţin **prin transformare lineară,** ele vor păstra relaţiile numerice exacte ale rezultatelor brute, deoarece sunt obţinute prin scăderea unei constante din flecare rezultat brut şi, apoi, împrăştierea rezultatelor conform altei constante. Într-un astfel de model, mărimea relativă a diferenţelor dintre rezultatele standard derivate va corespunde exact celei dintre rezultatele brute şi toate proprietăţile distribuţiei iniţiale a rezultatelor brute vor fi reproduse în distribuţia rezultatelor standard.

Tipuri de scoruri standard

1. Cotele Z

Rezultatele derivate linear se numesc şi **rezultate standard** sau **rezultate Z**. **Calcularea cotei Z** cere scăderea din rezultatul brut al persoanei a mediei grupului normativ şi împărţirea acestei diferenţe la abaterea standard a grupului. Orice rezultat brut egal cu **media** va avea valoarea lui **z = 0**. Abaterea standard este în cazul cotelor Z egală cu 1. Formula notelor Z este următoarea:

$$Z = \frac{x_i - \overline{\overline{x}}}{\sigma}$$

unde: x_i = scorul brut, $\overline{\overline{x}}$ = media scorurilor brute, σ = abaterea standard a scorurilor brute

Avantaj: între rezultatele Z avem aceeaşi distanţă.

Dezavantaj: în situaţia în care media e mai mare ca rezultatul, rezultatul e negativ.

2. Cotele de tip T şi **cotele de tip H**

Au intervenit pentru a facilita exprimarea rezultatelor sub medie, pentru a elimina rezulatele exprimate negativ.

In cotele T, se consideră că distribuţia are media 50 şi abaterea standard 10; în practică, scara T are limitele între 15 şi 8.

$$T = 50 + \frac{10\left(x_i - \overline{\overline{x}}\right)}{\sigma}$$

3. În cotele Hull,

Media este 50, abaterea standard 14, iar formula de calcul devine:

$$H = 50 + \frac{14\left(x_i - \overline{\overline{x}}\right)}{\sigma}$$

In mod artificial s-a introdus valoarea unei abateri standard de 14 în loc de 10. Pentru normare se calculează media, abaterea standard, apoi scorul brut se transformă în scor standard - cu un continuum standardizat ce se întinde de la 0 la valoarea maximă. Mai există şi alte tipuri de scoruri standard, însă cele trei sunt cele mai folosite.

Clasele normalizate

Un motiv principal al acestei proceduri constă în faptul că **cele mai multe distribuţii ale rezultatelor** brute, mai ales pentru teste de abilităţi, sunt aproape de curba normală a lui Gauss.
Rezultatele standard normalizate sunt rezultate standard exprimate în termenii unei distribuţii ce a fost transformată pentru a se potrivi curbei normale de distribuţie.
Repetăm proprietăţile curbei normale de distribuţie:

1. Această curbă are proprietăţi matematice care se pot constitui ca o curbă de rezultate statistice;
2. Scăderea rezultatelor se face pe un pas egal şi treptat spre cele două sensuri, pe măsură ce ne apropiem de extreme;
3. Simetria bilaterală - având un singur vârf la mijloc.

Există diferite modalităţi de a normaliza:

1. împărţirea acestei curbe în 5 clase normale;
2. împărţirea acestei curbe în 7 clase normale;
3. împărţirea acestei curbe în 9 clase standard (stanine);
4. împărţirea în 11 clase standardizate.

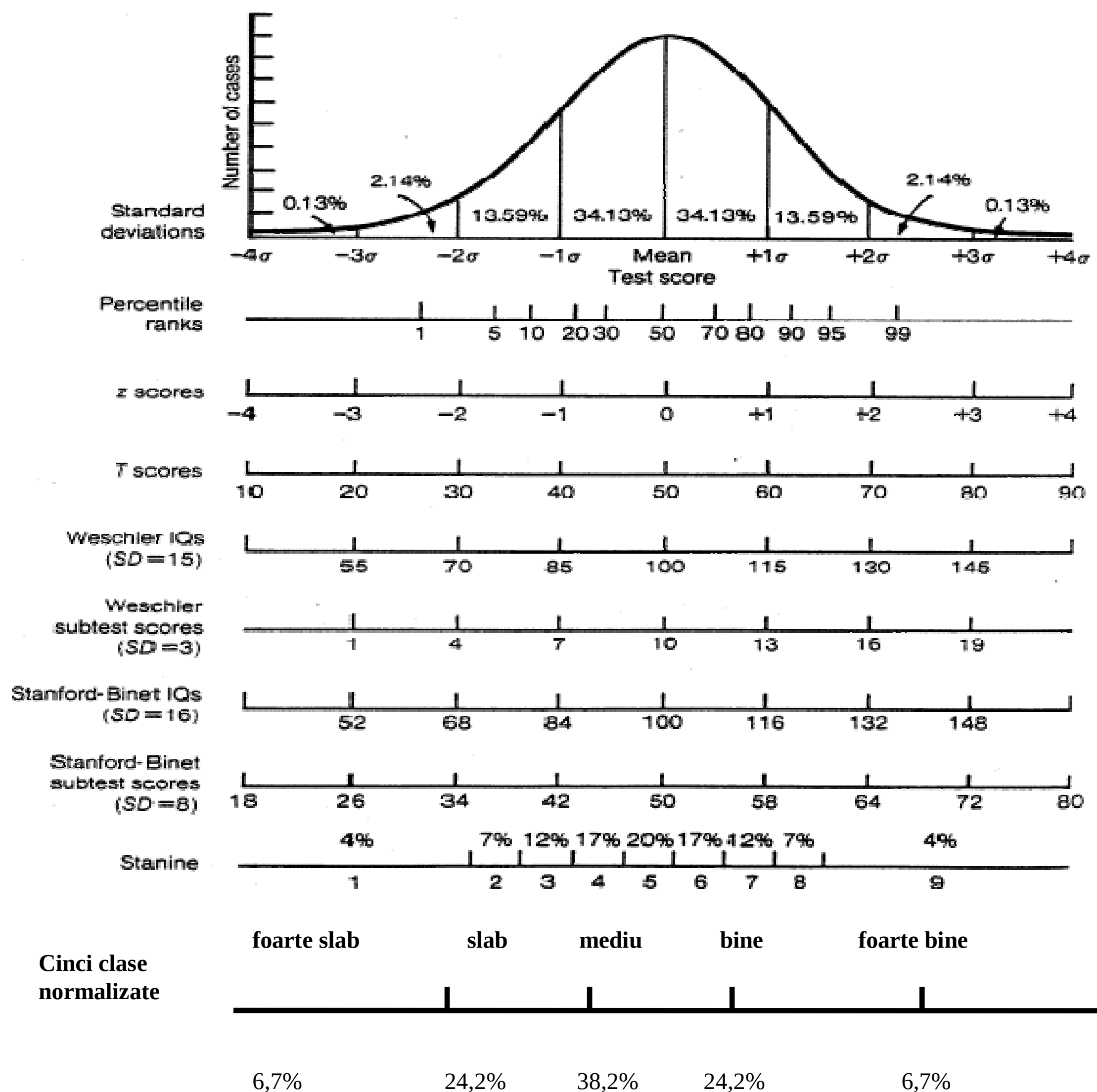
Paşii procedurii includ:
Calcularea tabelului de frecvenţe şi împărţirea lotului de subiecţi în unităţi procentuale egale, respectiv intervale între repere care nu sunt echidistante.

Procente pentru fiecare clasă

Etalonul constituit **5 clase normalizate** are ca procente: 1) 6.7%, 2) 24,2%, 3)38.2%, 4)24.2% şi 5)6.7% dintre subiecţii lotului de referinţă.
Etalonul în **7 clase normalizate:** 1)4.8%, 2)11.1%, 3)21.2%, 4)25,8%, 5)21,2%, 6)11,1% şi 7)4.8%
Etalonul în **9 clase (staninele)** are ca procente: 1)4.0%, 2)6.6%, 3)12.1%, 4)17.5%, 5)19.6%, 6)17.5%, 7)12.1%, 8)6.6%, 9)4.0%.
Etalonul în **11 clase (**de obicei util pentru testele de personalitate): 3.6%, 4.5%, 7.7%, 11.6%, 14.6%, 16.0%, 14.6,11.6%, 7.7%, 4.5%, 3.6%).

Acest tip de transformări nonlineare se efectuează numai când există un eşantion numeros şi reprezentativ şi când deviaţia standard de la rezultatele testelor se datorează defectelor testului şi nu caracteristicilor eşantionului sau altor factori care afectează eşantionul. Când distribuţia reală a rezultatelor brute se apropie de curba normală de distribuţie, rezultatele standard derivate linear şi rezultatele standard normalizate vor fi aproape identice. în astfel de situaţii, rezultatele standard şi clasele normalizate vor servi aceluiaşi scop. Reproducem mai jos echivalentele între diferitele tipuri

de scoruri la test:



Comparaţii a diferite tipuri de scoruri la test:

Tip de	Avantaj	Dezavantaj
scor brut	Dă un număr precis de puncte obţinute la test	Nu poale fi interpretat sau comparat
rangul percentil	Este mai uşor înţeles de utilizatori. Nu cere statistici sofisticate. Indică o poziţie relativă a scorurilor în percentile. Mai adecvat pentru datele fără distribuţie normală (distribuţii asimetrice).	Utilizează unităţi de măsurare ordinale; utilizează deci unităţi inegale de măsurare. Nu permite o tendinţă centrală. Nu poate fi comparat dacă grupele nu sunt similare. Distorsionează mult diferenţele de scor în partea de sus şi de jos a distribuţiei.
stanine	Scorurile sunt unitare. Poate fi exprimată tendinţa centrală. Simple şi utile.	Se poate ajunge să nu permită suficiente unităţi de scor pentru a putea să diferenţiezi printre scoruri. Sunt insensibile la diferenţele de mărime din cadrul staninei.
scoruri standard	Derivate din proprietăţile curbei normale. Reflectă diferenţele absolute dintre scoruri. Se poate calcula tendinţa centrală şi corelaţia. Dacă grupele de referinţă sunt echivalente, permit compararea de la test la test.	Inadecvate pentru datele ce nu se grupează conform curbei normale. Sunt dificil de explicat pentru utilizatori.

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice,* Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition),*Routledge, London and New York,
Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului,* Ed. Fundatiei Romania de Maine, Bucuresti,
Mitrofan Niculae- Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare,* Editura Polirom, Iaşi
Urbina Susana, *2004- Essential of psychological testing,* John Wiley & Sons, Inc. Hoboken, New Jersey

10. APLICAREA TESTULUI ŞI INTERPRETAREA REZULTATELOR

.

10.1. RELAŢIA TEST-OBIECTIVE INVESTIGATE. FACTORI CE INTERVIN ÎN TESTAREA PSIHOLOGICĂ

Testele psihologice evaluează fie atribute psihice, fie pot servi predicţiei unor comportamente. Testul constă din una sau mai multe probe care determină flecare o reacţie psihologică, comportamentală a subiectului sau îi cere să evalueze un comportament propriu sau al altuia. Răspunsurilor, conform unor reguli fixe, li se atribuie valori numerice, denumite scoruri.

Un item al testului, respectiv situaţia stimul care solicită un astfel de răspuns sau reacţie din partea subiectului, poate consta dintr-o întrebare, un desen, o problemă situaţională, o listă de alegere sau ordonare- răspunsului subiectului atribuindu-i-se un scor. Un set de itemi omogeni, care se referă la acelaşi aspect psihic, formează o scală. Pentru fiecare scală se calculează scorul pe baza itemi lor componenţă, conform unor reguli. Nu există restricţii privind numărul de itemi care formează o scală. Nu este necesar ca itemii unui test construit în scopul diagnosticului sau predicţiei unui comportament să descrie exact acel comportament.

De exemplu, pentru selecţia profesională ar fi imposibil ca un test să se refere la totalitatea situaţiilor şi comportamentelor profesionale specifice. Nu li se poate cere persoanelor ce se prezintă la examenul pentru permisul de conducere să conducă automobilul atâta vreme cât nu au urmat cursul de instruire. Testul psihologic măsoară comportamente prin care se pot face inferenţe asupra altor comportamente.

Taxonomia tipurilor de itemi realizată de A. Angleitner, 1986, se referă la relaţiile posibile între itemul testului de personalitate şi trăsătura evaluată. Astfel, itemul poate descrie 7 categorii de aspecte:

- reacţiile subiectului (deschise, comportamente observabile); reacţii acoperite, interne (precum sentimente, cogniţii interioare); reacţii psihofiziologice;
- atribute ale trăsăturii (dispoziţii, de obicei descrise prin adjective sau substantive);
- dorinţe şi trebuinţe (ca, de exemplu, intenţia de a se angaja în comportamente specificate);
- fapte biografice;
- atitudini (opinii puternic susţinute faţă de diverse categorii de subiecte generale, sociale, personale);
- reacţiile altora (reacţii, atitudini, opinii ale altora faţă de persoană);
- alte conţinuturi, uneori fără legătură aparentă.

Toate regulile privind caracteristicile psihometrice ale testului şi calităţile testului psihologic au ca obiectiv să menţină capacitatea testului de a măsura ceea ce îşi propune să măsoare.

Există o serie de **factori care intervin şi influenţează reacţia subiectului** la conţinutul testului şi care astfel pot interfera cu obiectivul acestuia, cu ceea ce îşi propune testul să măsoare:

I. **Factori ce ţin de caracteristicile stimulilor**
   a. Conţinutul testului
   b. Formatul testului: în relaţie cu viteza de lucru, maniera de răspuns, uşurinţa sau dificultatea de a face faţă acestui format
   c. Secvenţa sau modelul secvenţial în care se desfăşoară proba (paternul de succesiune a itemilor)
   d. Legătura culturală a caracteristicilor testului (dependenţa de cultura din care a făcut parte cel ce a făcut testul).

II. **Factori ce ţin de caracteristicile situaţiilor**
   a. Modul de administrare a testului: instructaj, timp de lucru standard, tipul de material cu care lucrează subiectul (standardizate)
   b. Contextul interpersonal, unde este inclusă şi influenţa examinatorului.
   c. Caracteristicile fizice: lumină, zgomot, stimuli luminoşi.
   d. Contextul social: este important dacă subiectul lucrează singur sau în grup, dacă metoda de înregistrare a rezultatelor ţine tot de caracteristicile situaţiilor.

III. **Factori ce ţin de caracteristicile persoanelor implicate**
   a. Scopurile personale ale testării („vreau să obţin un job sau vreau să le demonstrez celorlalţi că pot obţine postul").
   b. Stilurile personale de răspuns, care includ:
      II. dezirabilitatea socială (unii sunt mai legaţi de dezirabilitatea socială; în cadrul conţinutului va interveni acest al doilea factor, pe lângă
      conţinutul întrebării)
      III. complianţa (se referă la comportamentul de acceptare a ceea ce li se cere);
      IV. apărările sociale (ale eului): când, în structura personală, unul dintre mecanisme devine rigid, funcţionând în toate situaţiile de viaţă, atunci devine disfuncţional
   c. Caracteristici ce ţin de fiziologie: sănătate, oboseală, caracteristici de motricitate, auz etc.
   d. Fluctuaţii ale atenţiei şi memoriei, care pot ţine de o dinamică inconştientă.
   e. Capacitatea de a înţelege întrebările şi itemii testului şi îndemânarea în lectură (citire).
   f. Starea emoţională în sine şi starea de tensiune (calm rezultat din blocarea tensiunii - un hipercontrol).
   g. Măsura în care subiectul este familiarizat cu testarea.
   h. Dispoziţia de a avea sau nu noroc în ghicirea răspunsului (intuiţia este în rezonanţă profundă cu materialul).

i. Elemente care ţin de trăsăturile de comportament, motive, conflicte.

10.2 TESTAREA ŞI SURSELE VARIAŢIEI ERORII

Înainte de a calcula coeficientul de fidelitate al unui test este necesară precizarea surselor posibile ale erorilor de măsură. O caracteristică a unei persoane poate fi măsurată prin scorurile reale ale unui test sau poate influenţa scorurile observate, ca eroare de măsură, la un alt test. Condiţiile de administrare a testului (lumina, căldura, zgomotul etc), când diferă de cele standardizate, pot modifica destul de mult scorurile observate.

Din aceste motive, nu este suficient să se indice o valoare a coeficientului de fidelitate pentru a arăta cât de bine sunt reflectate scorurile reale de scorurile observate, ci trebuie să se precizeze:

- caracteristicile persoanelor cărora le este destinat testul (nivel de studii, vârstă etc);
- condiţiile în care se administrează testul.

Au fost detectate mai multe surse ale variabilităţii scorurilor, sintetizate în şase categorii.

Sursele posibile de variabilitate care pot afecta scorurile unui anumit test sunt:

1. particularităţile durabile şi generale ale persoanei;
2. particularităţile durabile, dar specifice ale persoanei;
3. particularităţi temporare, dar generale ale persoanei;
4. particularităţi temporare şi specifice ale persoanei;
5. factori sistematici sau de noroc care afectează administrarea testului sau evaluarea performanţelor la test; şi, în ultimă instanţă,
6. unele variaţii neexplicate precum şansa în găsirea răspunsului corect (subiectul ghiceşte) sau distragerea temporară a atenţiei (Thorndike, 1949; Murphy şi Davidshofer, 1991).

Particularităţile durabile, dar generale au influenţă asupra rezolvării mai multor teste. Ele pot fi: nivelul unor deprinderi ce pot influenţa nivelul de operare la teste; deprinderile şi tehnicile generale solicitate în testare; abilitatea generală de a înţelege instrucţiunile la teste.

Particularităţile durabile, dar specifice au influenţă numai asupra unui test. Ele pot fi specifice faţă de testul luat ca întreg, precum nivelul individual de dezvoltare a abilităţii respective care nu influenţează rezolvarea altor teste; cunoştinţe şi deprinderi specifice; modele de răspuns stabile. Pot fi, de asemenea, specifice pentru anumiţi itemi ai testului, precum: şansa persoanei de a şti răspunsul la un anume fapt particular cerut de item; tipuri de itemi cu care subiectul este anterior familiarizat.

Particularităţile temporare, dar generale ale persoanei se referă la factori incidentali care afectează performanţele nu numai la un anume test, ci la mai multe sau la toate testele, precum: sănătatea, oboseala, motivaţia, tensiunea emoţională, nivelul prea elevat al testului, înţelegerea sau nu a mecanismelor testării respective, condiţiile de administrare.

Particularităţile temporare şi specifice care influenţează doar un anume test pot fi specifice pentru testul în întregime sau faţă de anumiţi itemi ai săi. Din prima categorie fac parte: înţelegerea sarcinii specifice cerute de testul dat; micile trucuri sau tehnici de rezolvare a testului; nivelul de antrenare în rezolvarea unor anumite tipuri de sarcini; setul sau dispoziţia momentană. Din a doua categorie fac parte fluctuaţiile şi idiosincraziile caracteristice individului şi fluctuaţiile neprevăzute ale atenţiei sau percepţiei suprapuse nivelului general de performanţă caracteristic persoanei.

Dintre factorii sistematici amintim: condiţiile de testare (un anumit interval de timp fixat pentru testare, factori de distragerea atenţiei, claritatea instructajului); efectul interacţiunii personalităţii, sexului sau rasei examinatorului cu personalitatea celui examinat, care poate stimula sau inhiba performanţa; inconsecvenţa sau părtinirea în evaluarea performanţelor.

10.3. CERINŢE PRIVIND CONDUITA UTILIZATORULUI DE TESTE

Standardele de testare psihologică şi educaţională prevăd o serie de reguli privind utilizatorul de teste. Un principiu fundamental pentru examenul psihologic, indiferent de natura acestuia, este onestitatea ştiinţifică şi profesională a celui care examinează şi interpretează datele. In virtutea acestui principiu, este important ca în conduită să se evite complet atitudinea de „superioritate a-toate~cunoscătoare” care lasă să se înţeleagă că prin examinarea psihologică se ajunge să ştii totul despre persoana examinatului. Trebuie cunoscute şi avute mereu în vedere limitele procedurilor standardizate.

In examinare este necesară aplicarea probelor într-o manieră realistă, raţională şi favorabilă pentru subiectul examinat reducându-se la minim elementele care accentuează stresul în situaţia de examen.

Valoarea rezultatelor poate fi alterată semnificativ de lipsa de cooperare, de influenţa factorilor ce pot inhiba conduita normală a subiectului sau îl împiedică să realizeze nivelul de performanţă de care este în mod obişnuit capabil. Astfel de factori trebuie cunoscuţi de examinator şi înlăturaţi dintre variabilele situaţiei de examen, ca şi diferite atitudini ale subiectului care afectează modul său de lucru precum: teama de eşec, nevoia de a apărea într-o lumină favorabilă, percepţii deformate privind examenul, sensul acestuia şi examinatorul.

Printre condiţiile care favorizează eliminarea tensiunii şi care ţin de comportamentul examinatorului amintim: găsirea unei poziţii în care subiectului să-i fie mai puţin vizibile notaţiile sau observaţiile, sau aparatele de înregistrare a indicatorilor de eroare şi timp de execuţie; în general, pentru a scădea gradul de nesiguranţă, examinatorul trebuie să fie vizibil subiecţilor. În conduită se cer evitate descurajările prin afirmarea peremptorie a erorii, dar şi confirmarea continuă a corectitudinii răspunsurilor. Încurajarea, necesară în relaţia cu subiectul, nu trebuie să influenţeze direcţia şi calitatea gândirii şi conduitei subiectului. Se cer evitate manifestările de oboseală, plictiseală, nemulţumire.

Capacitatea de a face interpretări, de a înţelege natura dezadaptării sau insuccesului profesional sau existenţial, de a prognoza eficienţa într-o activitate cere psihologului o continuă perfecţionare prin cunoaşterea la zi a literaturii referitoare la

validitatea probelor respective şi lărgirea orizontului de cunoaştere profesională, aspecte care îi pot permite aprecierea corectă a metodelor la care poate recurge într-o situaţie dată, sau a acelor aspecte care trebuie adâncite în cadrul examinării.

În profilul psiho-comportamental este important să se manifeste calităţi precum: cultura psihologică, experienţa interpersonală câştigată inclusiv în contactul cu diferite categorii de persoane, autocunoaşterea, deprinderi formate în cercetarea psihologică şi/sau experimentală, dar şi calităţi psihice precum simţul empatic, capacitatea de a obţine cooperare, inteligenţa, sănătatea psihică. în mod specific sunt necesare memorarea corectă a instructajelor, deprinderea de a observa şi nota corect răspunsurile şi observaţiile.

10.4. RAPORTUL-CARACTERIZAREA FINALĂ

Forma pe care o ia acest raport este de obicei direct legată de tipul de examen şi de cerinţele beneficiarului. Raportul unui examen clinic diferă în amploare şi profunzime a interpretărilor de raportul unui examen de selecţie. Raportul privind consilierea educaţională diferă în conţinut şi tip de centrare pe probleme de raportul de expertiză judiciară. Rapoartele diferă, de asemenea, şi în funcţie de nivelul de comprehensibilitate al beneficiarului.Condiţiile şi principiile generale de întocmire a raportului sunt prezentate în *Standardele de testare educaţională şi psihologică.* Principiile respective se referă la obiectivitate, non-intruziunea, lipsa de ambiguitate şi subiectivism, respectarea confidenţialităţii, neutilizarea unor termeni care nu sunt inteligibili beneficiarului şi explicarea semnificaţiei rezultatelor prezente.

O deficienţă constă în descriptivismul exagerat (privind comportamentul subiectului în timpul examinării) fără a oferi şi comentarii asupra acestor prezentări. Simpla enumerare a observaţiilor făcute, fără comentarii, şi integrarea lor într-un sens anume, pot genera o varietate largă de interpretări ale beneficiarului, conducând cel mai adesea spre imagini greşite asupra subiectului, încărcate de subiectivismul celui care citeşte un astfel de raport. în raport se includ explicaţiile posibile, atenţionând când acestea nu sunt certe. în situaţia unor rezultate contradictorii este important ca psihologul să prezinte obiectiv constatările şi să încerce explicarea inconstantelor, incertitudinilor, discordanţelor dintre date.

Bibliografie minimală

Albu, Monica, 1998- *Construirea şi utilizarea testelor psihologice,* Editura Clusium, Cluj-Napoca,
Anastasi Anne, Urbina Susana, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,
Kline, P., 2000- *The handbook of psychological testing (Second Edtition),*Routledge, London and New York,
Minulescu Mihaela,2003- *Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului,* Ed. Fundatiei Romania de Maine, Bucuresti,
Mitrofan Niculae- Stan Aurel, 2007- *Testul psihologic. Evoluţie, construcţie, aplicare,* Editura Polirom, Iaşi

Urbina Susana, *2004- Essential of psychological testing*, John Wiley & Sons, Inc. Hoboken, New Jersey

Bibliografie curs

Albu, Monica, 1998, *Construirea şi utilizarea testelor psihologice,* Editura Clusium, Cluj-Napoca,
Albu, Monica, 2000, *Metode şi instrumente de evaluare în psihologie*, Editura Argonaut, Cluj-Napoca
Anne Anastasi, Susana Urbina, 1997- *Psychological Testing, (Seventh Edition),* Simon&Schuster, Upper Saddle River, New Jersey,
Cohen, R.J. et *al*. (1996). *Psychological Testing and Assessment. An Introduction to Tests and Measurement*. Mayfield Publishing Company, Mountain View, ed. a 3-a.
Cohen, R.J., Swerdlik, M.E. (2005). *Psychological Testing and Assessment*. *An Introduction to Tests and Assessment.* McGraw-Hill International Edition.
Creţu Romeo Zeno, Evaluarea personalităţii. Metode alternative, Editura Polirom, Iaşi
150 de experimente în psihologie pentru cunoaşterea celuilalt, Editura Polirom, Iaşi
Dafinoiu Ion, 2007, Personalitatea. Metode calitative de abordare: Observaţia şi interviul, Editura Polirom, Iaşi
Edenborough Robert, 1996- *Using psychometrics: A practical guide to testing and assessment,* Kogan Page Limited, London,
Edenborough Robert , 1996- *Effective interviewing: A handbook of skills, techniques and applications*, Kogan Page Limited, London
Edenborough Robert, 2007- *Assesment method in recruitment, selection & performance: a managers guide to psychometric testing, interviews and assessment centers,* Kogan Page Limited, London,
Feuerstein Reuven , Ya'Acov Rand, John E. Rynders, 1988- *Mediated learning- Teaching, task and tools to unlock cognitive potential,* Corwin Press, Sage Publications Company, Thousand Oaks, California,
Feuerstein Reuven, 1980- *Instrumental Enrichment*
Feuerstein Reuven, Ya'acov Rand, Mildred B Hoffman, 1979- *The dynamic assessment of retarded performers- The learning potential assessment device, theory, instruments, and techniques,* University Park Press, Baltimore,
*Mediated Learning In and Out of the Classroom (Mediated Learning in & Out of the Classroom),*by Cognitive Research Program (Author) University of Witwatersrand
Friedenberg, L. (1995), *Psychological Testing. Desigh, Analysis, and Use*, Allyn & Bacon.
Gary Groth Marnat, Aiken Lewis R., 2005- *Psychological testing and assessment,* John Wiley &Sons, inc.,
Gary Groth- Marnat, 2003- *Handbook of psychological assessment*, John Wiley &Sons, inc.
Gregory, R.J. (1996). *Psychological testing*. *History*, *Principles*, *and Applications*. Needham Heights: Allyn & Bacon.

Hăvâreanu Cornel, 2007, Cunoaşterea psihologică a persoanei. Posibilităţi de utilizare a computerului în psihologia aplicată, , Editura Polirom, Iaşi
Kaplan, M. Robert, Saccuzzo, P.D (2005), *Psychological Testing. Principles, Applications,And Issues*. Wadsworth: Thompson.
Kerliger Fred N., Howard B. Lee, 2006- *Foundations of behavioral research*, Harcourt College Publishers, California State University,
Kline, P., 2000-*The handbook of psychological testing (Second Edtition),*Routledge, London and New York,
Kulcsar Tiberiu, 1976, Lectii practice de psihodiagnoza, Universitatea Cluj- Napoca, Cluj Napoca
Marczyk Geoffrey R., DeMatteo David, and Festinger David S., 2004- *Essentials of Research Design and Methodology* , John Wiley&sons, inc., Hoboken, New Jersey,
Mariana Rosca, 1975, Metode de psihodiagnostic, Ed. Didactica si Pedagogica, Bucuresti
Matthews G., Psihologia personalităţii. Trăsături, cauze, consecinţe, Editura Polirom, Iaşi
Mandia Mentis, Marylin Dunn- Bernstein, Martene Mentis, 2008- *Don't Accept Me As I Am: Helping Retarded People to Excel,*
Mein Ben-Hur, 2008- *On Feuerstein"s Instrumental Enrichment- A collection*
Edited by, Sky Light, Training and publishing inc.
Minulescu Mihaela, 2003- Teorie si practica in psihodiagnoza. Fundamente in masurarea psihologica. Testarea intelectului, Ed. Fundatiei Romania de Maine, Bucuresti,
Minulescu Mihaela, 2001, Bazele psihodiagnosticului, Ed Universitatii Titu Maiorescu, Bucuresti
Minulescu Mihaela - 2002, Conceptul de normalitate psihica, in Dimensiunea psihosociala a practicii medicale, Ed. Infomedica, Bucuresti
Minulescu M., Iliescu, D., Nedelcea, C., 2005, *Chestionarele nonverbale de personalitate: NPQ si FF NPQ*, Editura PsihoCover, Bucuresti,
Minulescu Mihaela, 1996, Evaluarea psihologica in selectia profesionala, ed. Pan, Bucuresti
Minulescu Mihaela, 2007, Psihodiagnoza moderna. Chestionarele de personalitate, Ed. Fundatiei Romania de Maine, Bucuresti
Minulescu Mihaela, 1996, Chestionarele de personalitate in evaluarea psihologica, Garell Publishing House, Bucuresti
Mitrofan Niculae, 1997, Testarea psihologica a copilului mic, Ed. Mihaela, Bucuresti
Mitrofan Niculae, 2001, Psihometria si directiile ei de dezvoltare la inceput de milenium, in " Psihologia la inceput de millennium", Polirom, Iasi
Mitrofan Nicolae, Mitrofan Laurenţiu, 2008,Testarea psihologică. Inteligenţa şi aptitudinile, Editura Polirom, Iaşi
Muchielli Alex, 2007, Dicţionar al metodelor calitative în ştiinţele umane şi sociale, Editura Polirom, Iaşi
Nicola Grigore, 2008, Istoria psihologiei, Ed. Fundatiei Romania de Maine, Bucuresti
Pitariu Horia, Albu Monica, 1996, Psihologia personalului. Masurarea si interpretarea diferentelor individuale, Ed. Presa Universitara Clujeana, Cluj Napoca
Pitariu, H., Iliescu,D.,Tureanu,V., Peleasa - *CPI- Inventarul Psihologic California*,,C., Ed. Psiho Cover, Bucuresti, 2006
Radu, I. (coordonator), Miclea, M., Albu, M., Szamosközi, S., Moldovan, O., Nemeş, S., *Metodologie psihologică şi analiza datelor*, Editura Sincron, Cluj-Napoca, 1993.

Schiopu Ursula, 1976, Introducere in psihodiagnostic, Universitatea Bucuresti
Stan Aurel, 2007, Testul psihologic. Evoluţie, construcţie, aplicare, Editura Polirom, Iaşi
Ticu Constantin, 2008, Evaluarea psihologică a personalului, Editura Polirom, Iaşi
Urbina Susana,*2004- Essential of psychological testing* John Wiley & Sons, Inc. *Hoboken, New Jersey,*
Yin R., Studiul de caz. Design-ul, colectarea şi analiza datelor, Editura Polirom, Iaşi